Molecular Comparison of Virulent and Non-virulent *Alcelaphine Herpesvirus*-1 Derivatives

Jennifer Ann Handley
Moredun Research Institute,
Edinburgh, Scotland.

A thesis submitted for the degree of
Doctor of Philosophy
University of Edinburgh
1993

DECLARATION

The initiation, design and execution of the experimental work and the interpretation of the results presented in this thesis was carried out by the author. Contributions to the work of this thesis by colleagues is fully acknowledged in the text.

Jennifer Handley
October 1993

ACKNOWLEDGEMENTS

I would like to thank the following people for their valuable assistance in the preparation of this thesis:

Dr H.W. Reid for his enthusiasm, encouragement and guidance throughout this project, and for isolating the viral derivatives employed, in addition to performing all the animal work.

Dr D.R. Sargan for assisitance in analysis of the data, constructive advice on the content of this thesis, and for reading this thesis.

Dr A.J. Herring for advice on molecular biology and reading this thesis.

Dr N. Lally for practical advice on molecular biology, and for many helpful discussions.

Miss I. Pow for instruction, help and preparation of monolayered cell cultures.

Mr B.J. Easter and Mr P. Smith for all photographic work, and assistance in the preparation of this thesis.

Mr J. Spence, Mr G. Lindsay and Miss N. Flaherty for computing assistance.

Finally, I wish to thank my many colleagues in the laboratory for their help and support, especially Mrs S. Nicoll, Dr A.G. Rae, Mr I. Anderson, Mr K. McLean and Mrs S. Dunbar.

This work was supported by a flexibility funding award from the Scottish Office Agriculture and Fisheries Department.

ABSTRACT

Malignant catarrhal fever (MCF) is a fatal lymphoproliferative disease affecting ruminants. The disease is caused by infection of susceptible hosts with one of two gammaherpesviruses, *Alcelaphine Herpesvirus*-1 (AHV-1) or *Ovine Herpesvirus*-2 (OHV-2).

On isolation AHV-1 infectivity is cell-associated and will induce MCF on inoculation into experimental animals. After serial passage cell-free infective virus is observed, but this virus cannot produce MCF experimentally. The aim of this project was to characterise the genomic alterations which occurred in one isolate, C500, associated with this altered pathogenicity.

Viral DNA of virulent (PP) and cell-free attenuated (CFA) C500 derivatives was compared by restriction endonuclease profiling. Variability was observed using Sma I, as a 5kbp fragment was present in PP DNA but not CFA DNA. Conversely a 3.8kbp fragment was present in CFA DNA but absent from PP DNA. Homology was observed between these Sma I fragments. The Sma I 3.8kbp fragment was cloned (ATT-1), as were two Hind III equivalents of the PP 5kbp Sma I fragment (VIR-1 and VIR-2). These clones were mapped for several restriction endonucleases, subcloned and sequenced.

The location of the C500 clones was assessed by Southern blotting and PCR. The structure of the C500 genome consisted of a central unique region of approximately 130kbp, flanked on either side by multiple copies of a 1050bp repeat unit, resulting in an overall genome size of 160kbp. The ATT-1 clone was found to be present twice in the CFA genome, located at both ends of the unique region, close to the terminal repeats, with both copies orientated in the same direction. The location of the VIR-1 and VIR-2 clones in the PP genome was not fully determined during this study, however, the results obtained suggested that the VIR-1 and VIR-2 clones were located approximately 1.4kbp apart, at one end of the unique region, close to the terminal repeats. The mapping data suggested that two sequential deletions occur as the C500 virus is passaged in tissue culture, resulting in a total deletion of approximately 3.8kbp. This deletion correlates with loss of virulence.

The sequence of the three C500 clones was compared to other known sequences at both nucleic acid and amino acid levels, but no significant homology was

observed. One open reading frame (ORF), encoding a peptide designated peptide 5, and an adjacent unsequenced region of 800bp were lost on attenuation of the C500 isolate, suggesting that this region encodes sequence essential to the development of MCF. The sequence of peptide 5 did not exhibit structural similarity to the peptides involved in the transforming potential of either of the related gammaherpesviruses *Herpesvirus saimiri* (HVS) or Epstein-Barr virus (EBV). However, the location of the ORF encoding peptide 5 in the C500 genome was similar to that of the transforming protein of HVS in the HVS genome

One ORF, encoding a peptide designated peptide 1, was found to be conserved in OHV-2, and is a further indication of the close relationship between these two herpesviruses. Peptide 1 also exhibited limited structural homology to the transforming proteins of HVS, however, the presence of two copies of peptide 1 in the CFA C500 genome suggested that peptide 1 did not have a similar function to the HVS peptides.

TABLE OF CONTENTS

Chapter 1 - General Introduction

Chapter 2 - Materials and methods

Chapter 3 - Identification and characterisation of variable regions of C500

ABBREVIATIONS

Ab	antibodies
Ag	antigens
AHV-1	*Alcelaphine Herpesvirus*-1
AHV-2	*Alcelaphine Herpesvirus*-2
APS	Ammonium persulphate
BHV-4	*Bovine Herpesvirus*-4
bp	base pairs
CA	cell-associated virulent C500 virus
CF	cell-free virulent C500 virus
CFA	cell-free attenuated C500 virus
CIP	calf intestinal phosphatase
cpe	cytopathic effects
DIG	digoxygenin
DIP	defective interfering particles
DNA	deoxyribonucleic acid
ddNTP	di-deoxynucleotide triphosphate
dNTP	deoxynucleotide triphosphate
DMSO	dimethyl sulphoxide
EBNA	Epstein-Barr virus nuclear antigen
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetra-acetic acid
FBS	foetal bovine serum
GCG	genetic computer group
HSV	Herpes simplex virus
HSUR	*Herpsevirus saimiri* U RNA
HVA	*Herpesvirus ateles*
HVS	*Herpesvirus saimiri*
IgG	Immunoglobulin G
IIF	indirect immunofluorescence
IL-2	interleukin-2
IMDM	Iscove's modification of Dulbecco's medium
IPTG	isopropyl-D-thiogalactopyranoside
LAK	lymphokine activated killer cells
L-amp	L-ampicillin
L-broth	Lauria broth
LGL	large granular lymphocytes
LMP	latent membrane protein
MCF	malignant catarrhal fever
mRNA	messanger RNA
NK	natural killer cells
NP40	nonidet P40
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

Chapter 1
General Introduction

1.1 INTRODUCTION

Malignant catarrhal fever (MCF) is a fatal lymphoproliferative disease affecting cattle and several other ruminant species. The disease has a worldwide distribution, but its incidence is usually sporadic and low, with no particular reference to age, sex or breed (Plowright 1968). Epizootics have been observed in cattle, red deer (*Cervus elaphus*) and bison (Pierson *et al* 1973, Plowright *et al* 1975, Ruth *et al* 1977, Reid *et al* 1979).

MCF was first recognised in Europe, and later in Africa, where substantial losses of cattle were reported following contact with wildebeest (reviewed by Heuschele *et al* 1984). There have been reported cases of MCF in at least 39 species of ruminant worldwide (reviewed by Heuschele *et al* 1988). In Asia MCF is economically important as water buffalo (*Bubalus bubalis*) and Bali cattle/Banteng (*Bos javanicus* synonym *Bos sandaicus*) are particularly susceptible (Hatkin 1980, Daniels *et al* 1988). With the development of deer farming in the UK, Australia and New Zealand it has become evident that deer are especially susceptible to MCF (Reid *et al* 1979, Denholm and Westbury 1982, Oliver *et al* 1988).

The aetiology and pathogenesis of the disease is complex, with at least two viral agents capable of inducing MCF. In Africa the disease is caused by infection with the gammaherpesvirus *Alcelaphine Herpesvirus*-1 (AHV-1), which was first isolated from an apparently clinically normal blue wildebeest (*Connochaetes taurinus*) in 1959 (Plowright *et al* 1960). Outside Africa MCF is normally induced by infection with *Ovine Herpesvirus*-2 (OHV-2), which has not been isolated nor fully characterised, but domestic sheep are the natural, unaffected host of this virus (Rossiter 1981a, Baxter *et al* 1993). This virus is antigenically and genomically closely related to AHV-1.

1.2 CLINICAL FEATURES

The clinical features of MCF exhibited following infection of susceptible hosts with either aetiological agent are generally similar. The symptoms are, however, variable with the disease being categorised into "head and eye", peracute, intestinal and mild forms. The majority of cases in cattle are of the "head and eye" form, while the peracute form is more frequent in deer (reviewed by Plowright 1986).

Clinical symptoms are generally observed by sudden onset of persistant pyrexia, where the temperature of the affected animal typically rises to 42°C (Plowright 1986). Severe congestion, necrosis and erosion of the nasal and oral mucosae are recorded in most cases of AHV-1 induced disease (reviewed by Plowright 1986). In OHV-2 induced disease these symptoms are less frequently demonstrated, and when observed are generally less severe (James *et al* 1975, Reid *et al* 1979, Denholm and Westbury 1982). A generalised enlargement of lymph nodes is evident in animals affected with MCF. Further symptoms include diarrhoea, dysentery, dermatitis and laminitis, all of which tend to occur more frequently in cattle and deer affected with OHV-2 induced disease (Mare 1977, Reid *et al* 1979, Denholm and Westbury 1982). Signs of central nervous system involvement such as muscular tremours and lack of co-ordination have been observed in animals infected with AHV-1 and OHV-2 (Plowright 1986).

In deer the disease is often peracute and clinical manifestations can be minimal prior to death. The development of clinical features such as conjunctivitis, necrosis of the muzzle and dry necrotic skin lesions may appear in acute and chronic cases in deer, depending on the length of survival (Reid *et al* 1984).

MCF has been observed to develop as a result of experimental infection of rabbits using both aetiological agents (Daubney and Hudson 1936, Buxton and Reid 1980). The symptoms noted in this experimental model are largely similar to those exhibited in hosts naturally affected with the disease. A rectal temperature of greater than 40°C was generally the first evident clinical sign, followed by the rabbits becoming dull and anorexic. Serous or catarrhal nasal and ocular discharges developed, accompanied by conjunctivitis. The submandibular and mesenteric lymph nodes were often enlarged. Lymphoid accumulations were often observed in the liver, lungs, stomach, intestine, periphery of the cornea, and mildly in the kidneys, heart, reproductive organs, brain and spinal cord. In addition haemorrhages and necrotic foci frequently affected the liver, lungs and mucosal surfaces (Buxton and Reid 1984).

Incubation times between contact with the viruses and the onset of clinical symptoms have not been established as there is no method of determining precisely when an affected animal contacted the virus. Pierson *et al* (1973) suggested that the incubation time of OHV-2 in susceptible hosts varied between a minimum of 60 to a maximum of 144 days.

established on the overwhelming majority of properties where multiple cases occured (reviewed by Plowright 1986). In addition several publications have reiterated that infection of cattle and deer is most likely during the lambing season (Reid and Buxton 1984 and 1985, Reid *et al* 1984) and the season of highest incidence for cattle, often reported as late winter and spring to early summer in temperate zones, would be consistent with this hypothesis (Plowright 1986). Rossiter (1981a) further implicated sheep as viral hosts by demonstrating that a high proportion of sheep sera tested positive for antibodies to AHV-1 by an indirect immunofluoresence test (Rossiter 1981b). These antibodies were detected in 162 of 167 sheep sera collected from 10 different flocks (Rossiter 1981a). In addition an OHV-2 specific polymerase chain reaction (PCR) demonstrated OHV-2 in peripheral blood lymphocytes of 6/6 normal domestic sheep (Baxter *et al* 1993).

1.3.3 Related gammaherpesviruses of bovidae

In addition to wildebeest AHV-1 neutralising antibodies have been found in several other species of the subfamily *Alcelaphinae.* These include the hartebeest (*Alcelaphus buselaphus*) and at least 3 species in the genus Damaliscus (topi, tsessebe and blesbok), (reviewed by Plowright 1986). In addition to these neutralising antibodies, herpesviruses were isolated from free living hartebeest and topi (*Damaliscus korrigum*), (Mushi *et al* 1981), while in the USA herpesvirus was isolated from captive Cape hartebeest (*Alcelaphus buselaphus camaa*) and Jimela topi (*Damaliscus lunatus jimela*), (Heuschele *et al* 1984). A herpesvirus isolated from the hartebeest by Reid and Rowe (1973) was later classified as *Alcelaphine Herpesvirus*-2 (AHV-2), (Roizman *et al* 1981).

In the subfamily *Hippotraginae* there is also widespread evidence of infection with AHV-1-related herpesviruses in species of addax and oryx. Virus neutralising antibodies were found in fringe-eared oryx (*Oryx gazella biesa*) sera (Mushi and Karstad 1981), and an AHV-1-related agent was isolated from a healthy scimitar-horned oryx (*Oryx gazella dammah*), (Heuschele *et al* 1984). A herpesvirus antigenically and genomically related to AHV-1 has also been isolated from a roan antelope (*Hippotragus equinus*) and designated *Hippotragine herpesvirus*-1 (Reid and Bridgen 1991).

In addition to sheep, the presence of AHV-1 related viruses has been implicated in other members of the subfamily *Caprinae.* Neutralising antibody has been detected

in several species of goat, including the domestic goat (*Capra bircus*), and in both the alpine and nubian ibex (*Capra ibex ibex* and *Capra ibex nubiana*), (Heuschele *et al* 1984).

The gammaherpesvirus *Bovine Herpesvirus*-4 (BHV-4) exhibited a low, but consistent, level of serological cross-reactivity with AHV-1 (Osorio *et al* 1985).

Although the additional carriers described above harbour herpesviruses related to AHV-1, they can apparently be grazed or housed with species susceptible to MCF with no detrimental effect. As yet, only the wildebeest and the domestic sheep can be incriminated as a source of herpesviruses which result in MCF in susceptible hosts.

1.4 HOSTS SUSCEPTIBLE TO MCF

All breeds of cattle and domestic buffalo are at all ages susceptible to MCF. Judging from the frequency of cases reported in zoological collections, the species most susceptible to AHV-1 disease are the Asiatic cattle, banteng (*Bos javanicus*) and Indian Guar (*B. gaurus*). Some African antelope are also particularly susceptible to AHV-1 induced MCF such as the kudu (*Tragelaphus stresiceros*) and the sitatunga (*T. spekei*), (Plowright 1986).

The Cervidae appear to be much more susceptible than cattle to OHV-2 infection, as severe outbreaks of OHV-2 induced MCF affecting axis deer (*Cervus axis*), barasingha deer (*C. duvauceli*), red deer (*C. elaphus*), sika deer (*C. nippon*), Rusa deer (*C. timorensis*), roe deer (*Capreolus capreolus*) and Pere David's deer (*Elaphanus davidianus*) have been reported (Reid and Buxton 1984, Plowright 1986).

The remainder of the 39 species which demonstrate susceptibility to MCF are reviewed by Heuschele *et al* (1988).

Experimentally rabbits, hamsters, guinea-pigs, and rats have demonstrated susceptibility to AHV-1 and OHV-2 (Daubney and Hudson 1936, Buxton and Reid 1980, Jacoby *et al* 1988b). Experimental transmission is discussed further in section 1.6.

1.5 NATURAL TRANSMISSION

The mode of natural transmission from the reservoir host to the susceptible ruminant is probably via nasal and ocular secretions in the form of aerosols. This theory is substantiated by the observation that AHV-1 can be transmitted to rabbits by administering cell-free (CF) virus intranasally (Mushi and Rurangirwa 1981) and that CF AHV-1 virus has been isolated from the nasal and ocular secretions of wildebeest calves which were less than three months old (Mushi *et al* 1980a). In the case of OHV-2, since the virus has not yet been isolated, the method of transmission has not been elucidated, but is expected to be similar to that of AHV-1.

Studies to compare the stability of cell-associated (CA) and CF AHV-1 in the environment revealed that CA virus was very fragile, but CF virus could survive outwith the host for more than 13 days under conditions of high humidity (Mushi and Rurangirwa 1981). When infected with AHV-1, cattle act as a dead end host for the virus. Studies illustrated that, although cattle shed AHV-1 in nasal and ocular secretions, all infectivity was cell-associated. This suggested that the inability of AHV-1 to spread by contact between cattle is probably due to the absence of CF virus in secretions (Mushi and Rurangirwa 1981).

1.6 EXPERIMENTAL TRANSMISSION

AHV-1 infection has been transmitted to rabbits using brain tissue or blood from affected cattle as inoculum (Daubney and Hudson 1936). Experimental transmission of AHV-1 from cattle to deer using whole blood from infected cattle has also been achieved (Whitenack *et al* 1981). Successful transmission to neonatal hamsters, rats and guinea-pigs by inoculation of rabbit lymphoid cells infected with AHV-1 has been accomplished, however, with weaned animals only guinea-pigs were susceptible (Jacoby *et al* 1988a and b). Transmission to sheep was observed as a result of inoculation with blood from MCF affected cattle, but clinical and pathological signs of MCF were not observed (Kalunda *et al* 1981b).

Experimental transmission of OHV-2 MCF has been more erratic and difficult to achieve than that of AHV-1, requiring at least eight to ten times more blood as inoculum than transmission of AHV-1 (Pierson *et al* 1979). OHV-2 induced MCF has been transmitted, using blood from the affected animal as inoculum, from cattle

Liggitt and De Martini (1980b) suggested that the pathogenic mechanism of MCF may be analagous to a "graft versus host" reaction. Hunt and Billups (1979) proposed viral-induced lymphoid cell transformation on the basis that the primary feature of the infection is generalised lymphocytic proliferation and infiltration of most organs and tissues, and also because the histopathological features of MCF are similar to those induced by known oncogenic herpesviruses (e.g. *Herpesvirus saimiri*). Denholm and Westbury (1982) subscribed to this theory, while also suggesting an alternative unsupported hypothesis involving uncontrolled proliferation of T lymphocytes due to viral destruction or inactivation of T suppressor lymphocytes. Mushi *et al* (1980b) examined the idea that the effects of MCF are due to immuno-compromisation by treating infected rabbits with levamisole (a potent anthelmintic which restores T lymphocyte function), but this did not affect the outcome of the disease.

The hypothesis which satisfies most of the features of MCF was proposed by Reid and Buxton (1985), who suggested that a profound virus-induced immune regulatory dysfunction could be the fundamental lesion. The T lymphocyte hyperplasia was found to be a side-effect of immune deregulation rather than a cause of death since treatment of rabbits with cyclosporin A (a potent T lymphocyte suppressor) resulted in abrogation of the hyperplasia but did not prevent death in rabbits (Buxton *et al* 1984). Cells cultured from a cow, a deer and an experimentally infected rabbit (all exhibiting signs of OHV-2 induced MCF) were characterised as large granular lymphocytes (LGL) with natural killer (NK) cell properties. It was found that as few as 100 of these cultured rabbit cells could transmit MCF, implying that a high proportion of the cultured cells carried the agent of MCF (Reid *et al* 1985). Reid and Buxton (1985) further proposed that interleukin-2 (IL-2), or a similar cytokine, is produced by infected LGLs, and this causes a benign and probably polyclonal T lymphocyte hyperplasia, possibly aided by malfunction of the T lymphocyte suppressor role that LGLs are recognised to possess (Biron and Welsh 1982). Finally, the tissue damage may arise from unrestricted NK activity of infected LGL, or what should more appropriately be termed lymphokine activated killer (LAK) activity, causing cytolysis of normal cells.

To investigate further the role of these LGL in the pathogenesis of MCF a lymphoblastoid cell-line was derived following an attempt at fusion of lymph node cells from a rabbit clinically affected with OHV-2 induced MCF to cells of ovine origin. The resultant lymphoblastoid cell-line resembled LGL morphologically,

carried OHV-2 and transmission of MCF was achieved using this cell-line as inoculum (Reid *et al* 1983). Similar cell-lines were subsequently derived from affected cattle and deer (Reid *et al* 1989a) and characterised using a panel of monoclonal Ab to bovine and ovine surface Ag which are also cross-reactive with deer. Results showed that all lymphoblastoid cell-lines were T lymphocytes with 3/4 cattle lines $CD4^-/CD8^+/T19^-$, 1/4 $CD4^+/CD8^-/T19^-$ and 2/2 deer lines $CD4^-/CD8^-/T19^+$ (Burrells and Reid 1991).

1.8 ANTIBODY TO AHV-1

Evidence of infection with AHV-1, OHV-2 and related viruses in other hosts has depended largely on the detection of antibody reactive with AHV-1. Plowright (1967) demonstrated that virus neutralising antibodies (VN Ab) to AHV-1 were present in all free-living wildebeest in East Africa. Various methods to detect Ab to AHV-1 were attempted (Plowright 1968, Rossiter *et al* 1977, and 1978, Rossiter and Jesset 1980, Rossiter 1981b), with an indirect immunofluorescence (IIF) test being the most successful (Rossiter 1981a). Sheep sera were examined using this IIF test, and 162 out of 167 sheep from 10 different flocks showed positive results, suggesting that the majority of sheep are infected with an agent antigenically related to AHV-1 (Rossiter 1981a). Immunoblotting analysis, using the AHV-1 strain WC11 (Plowright *et al* 1965), showed that sheep sera recognised some of the same epitopes as wildebeest sera (Herring *et al* 1989).

1.9 MOLECULAR DETECTION OF AHV-1 AND RELATED VIRUSES

In recent years the detection of both forms of MCF has focussed on direct detection of virus rather than serological methods. Hsu *et al* (1990a) developed a method for detection of AHV-1 using the polymerase chain reaction (PCR) which successfully detected viral DNA from as few as 10 infected tissue culture cells (corresponding to picogram amounts of AHV-1 DNA). In addition, Katz *et al* (1991) developed a nested PCR which detected both AHV-1 and AHV-2. The feasibility of employing these tests as diagnostic aids has not, however, been appraised.

The detection of OHV-2 in domestic sheep and clinical cases of OHV-2 induced MCF has been achieved using PCR (Baxter *et al* 1993). The sequence used to generate the oligonucleotides used in this PCR was OHV-2 specific. Briefly, a cytotoxic lymphoblastoid cell-line was developed by culturing cells from an animal

affected with OHV-2 induced MCF on feeder monolayers (Reid *et al* 1983 and 1989a). A genomic library was prepared from this cell-line and was probed with AHV-1 fragments (Bridgen and Reid 1991). A clone which exhibited partial homology to AHV-1 was characterised and sequence unique to OHV-2 was identified. This sequence was used to generate the OHV-2 specific oligonucleotides employed in a sensitive PCR to detect OHV-2.

1.10 IMMUNITY AND ATTEMPTS TO IMMUNISE SUSCEPTIBLE HOSTS

1.10.1 The immune response of reservoir hosts.

This section refers almost exclusively to infection of the wildebeest with AHV-1. Comparative studies of OHV-2 have not been achieved due to the inability to isolate the OHV-2 virus.

A study of free-living wildebeest showed that the wildebeest calf is infected with AHV-1 either *in utero* or soon after birth by natural transmission (Plowright 1965a). The transmission of high titre maternal Ab via the colostrum does not protect the calves against active infection and viraemia occurs (Plowright 1967). Viraemia is more common in calves than in adults, and among adults detectable viraemia is most common in pregnant animals, probably because the virus becomes latent after initial infection and is only reactivated under conditions of exceptional stress and pregnancy (Plowright 1965a and b). The finding that betamethasone, an immunosuppressive corticosteroid, led to recrudesence of viraemia in wildebeest further confirms latency of the virus in this host (Rweyemamu *et al* 1974).

1.10.2 Immunity of susceptible hosts.

OHV-2 and AHV-1 induced MCF result in death of susceptible hosts in more than 95% of cases. On the rare occasions where cattle have been observed to survive infection with AHV-1 the recovered animals acquired a solid resistance to parenteral challenge (Plowright 1968). One particular cow which recovered from MCF subsequently passed infection congenitally to four calves. MCF was observed in two of these calves, and viraemia was observed in the other two. The mother did not develop clinical signs of MCF during pregnancy (Plowright *et al* 1972). This observation suggested that the animal was latently infected, with recrudescence occurring in pregnancy in a similar manner to that observed in wildebeest. Latency in recovered animals was also reported by Rweyemamu *et al* (1976).

1.10.3 Attempts to immunise susceptible hosts.

Plowright *et al* (1975) attempted to protect cattle using inactivated AHV-1 virus in various forms. Formalised Ag, combined with Freunds incomplete adjuvant regularly induced a good neutralising antibody response, but this was not found to be protective, as vaccinated animals developed MCF on parenteral challenge with virulent virus. Plowright suggested that subsequent vaccination studies should attempt to stimulate cell-mediated rather than humoral immunity (Plowright *et al* 1975).

Edington and Plowright (1980) demonstrated immunity to cell-free but not cell-associated challenge in rabbits. The vaccine used consisted of cell-free AHV-1 virus inactivated with formalin or acetylethylamine and combined with Frends complete adjuvant. Russell (1980) also reported immunity to cell-free AHV-1 in 4/6 rabbits vaccinated with inactivated, cell-free virulent virus. Rossiter developed a complex protocol which protected against cell-associated virulent virus in 4/4 rabbits following initial challenge, and 2/4 when rechallenged 43 weeks later (Rossiter 1982a).

The AHV-1 related virus isolated from hartebeest, AHV-2, was used to attempt to immunise against AHV-1. An isolate of AHV-2, K30, became attenuated after serial passage in tissue culture. This attenuated variant of K30 was used to inoculate 6 cattle. All 6 cattle developed serum neutralising Ab to the AHV-1 laboratory isolate WC11 but the cattle were not protected from AHV-1 challenge. The vaccination procedure was partially successful for AHV-2 as 2 of the 3 challenged with virulent K30 survived (Reid and Rowe 1973).

In summary, there have been several attempts to immunise susceptible hosts, some of which have been moderately successful but there is at present no vaccination programme. The observation that subclincally infected animals may become carriers and transmit virus to their offspring in a method similar to that of wildebeest congenital transmission must be considered in attempts to develop a vaccine.

There is no standard treatment for animals with MCF, therefore, in the absence of vaccines, the prevention of MCF relies on separation of susceptible animals from transmitter animals.

1.11 CLASSIFICATION OF AHV-1

1.11.1 Introduction

The virus which induces MCF in Africa was first isolated from a blue wildebeest in 1959 (Plowright *et al* 1960). In this initial paper the virus was described as producing a cytopathology reminiscent of herpes simplex and varicella zoster in monolayer cell cultures. Affinity of this wildebeest-derived virus to the herpes group was further implied with respect to its size, morphology, relationship to the host cell and its sensitivity to ether and chloroform (Plowright *et al* 1965). This virus was subsequently designated *Alcelaphine Herpesvirus*-1 (AHV-1), (Reid et al 1975).

Ovine Herpesvirus (OHV-2) has not been isolated in culture in a similar manner to AHV-1. Viral cytopathic effects have not been detected in cell cultures derived from affected tissues of rabbits or deer, nor in bovine or ovine cells co-cultivated with such cells (Reid *et al* 1979, Westbury and Denholm 1982). Various viruses have been identified in animals affected with OHV-2 induced MCF, but none of these viruses has reproduced MCF on reinoculation into cattle, suggesting that such viral infections have been fortuitous and have no causal relationship with the disease (reviewed by Reid *et al* 1984). A cytotoxic T-lymphocyte cell-line derived from a rabbit affected with experimentally induced MCF had characteristics similar to cell-lines generated from cotton top marmosets infected with *Herpesvirus ateles* and *Herpesvirus saimiri* (Reid *et al* 1983). A subsequent report described additional lymphoblastoid cell-lines derived from OHV-2 affected cattle and deer (Reid *et al* 1989a). DNA was extracted from these lymphoblastoid cell-lines, and homology to AHV-1 was observed (Bridgen 1991, Bridgen and Reid 1991). This homology, coupled with the serological data (Rossiter 1981a, Herring *et al* 1989, Reid *et al* 1989b) led to OHV-2 being classified as a herpesvirus (Roizman *et al* 1992).

The family Herpesviridae has been extensively studied and subsequently subdivided both in terms of biological criteria and molecular structure. The remainder of this section describes the properties of the family *Herpesviridae*, with reference to AHV-1.

1.11.2 Definition of a herpesvirus

Membership of the family *Herpesviridae* is based on the architecture of the virion. Herpes virions vary in size from 120-300nm in diameter and consist of a core

containing linear, double stranded DNA, an icosahedral capsid surrounded by an amorphous region known as the tegument, and an envelope containing viral spikes (Roizman 1990). Genomic analysis has confirmed that AHV-1 has a double stranded DNA core (Bridgen *et al* 1989), while electron microscopic observations showed extracellular virions of enveloped and non-enveloped particles ranging in size from 140 to 220nm. A viral tegument was also observed around an inner icosahedral capsid of approximately 100nm (summarised by Seal *et al* 1988). Thus morphologically AHV-1 is a candidate for inclusion in the herpesvirus family.

1.11.3 Nomenclature of herpesviruses

The herpesvirus study group of the International Committee for the Taxonomy of Viruses (ICTV) recommended a binomial system of nomenclature which they summarised to three main points: (i) the first name should reflect the family (when the name should end in -id) or the subfamily (when it should end in -ine) of the primary host; (ii) the second name should be herpesvirus and (iii) different herpesviruses with the same primary host should be given sequential arabic numbers (Roizman 1982).

The wildebeest is a member of the family *bovidae*, subfamily *alcelaphinae*. AHV-1 has also been referred to as *Bovid herpesvirus*-3 in some literature (Ludwig 1983), however since *Bovidae* is such a large and complex family Reid *et al* (1975) suggested the usage of the subfamily name *Alcelaphinae* to designate the virus which causes wildebeest-derived MCF. This nomenclature was proposed since AHV-1 was first isolated from a wildebeest (Plowright *et al* 1960) and also since a similar virus had been isolated from a hartebeest (Reid and Rowe 1973), another member of the subfamily *Alcelaphinae*. This hartebeest-derived virus, distinct from AHV-1, has been designated *Alcelaphine Herpesvirus*-2 (AHV-2), (Reid *et al* 1975, Roizman *et al* 1981). The sheep-associated agent was originally designated *Caprine herpesvirus-3* (CHV-3), since domestic sheep belong to the subfamily *Caprinae* (Bridgen and Reid 1991). However, to avoid confusion the virus was re-designated *Ovine Herpesvirus* 2 (Roizman *et al* 1992).

1.11.4 Biological properties of herpesviruses

The known herpesviruses appear to share four significant biological properties which are as follows: (i) all herpesviruses specify a large array of enzymes involved in nucleic acid metabolism and possibly processing of proteins; (ii) both the synthesis of viral DNAs and the assembly of capsids occur in the nucleus; (iii) production of

infectious progeny virus is invariably accompanied by the irreversible destruction of the infected cell and (iv) herpesviruses are able to remain latent in their natural hosts. Although the herpesviruses share these properties, they also vary greatly in their biological properties with respect to host range, efficiency of multiplication, the cell type in which they remain latent and the clinical manifestations of the disease they cause (Roizman 1990).

The biological criteria used for the original classification of herpesviruses into subfamilies were based on simple measurements or observations that could be readily determined in many laboratories (Roizman *et al* 1981). The herpesviruses were divided into three subfamilies which are described below.

Alphaherpesvirinae

The members of this subfamily are classified on the basis of a variable host range, a relatively short reproductive cycle, rapid spread in culture, efficient destruction of infected cells and the capacity to establish latent infections primarily, but not exclusively, in sensory ganglia (summarised by Roizman 1990). An example of an alphaherpesvirus is human herpesvirus-1 (herpes simplex virus-1, HSV-1).

Betaherpesvirinae

A non-exclusive characteristic of the members of this subfamily is a restricted host range. The reproductive cycle is long, the infection progresses slowly in culture with the infected cells frequently becoming enlarged (cytomegalia) and the virus can establish latency in secretory glands, kidneys and other tissues (summarised by Roizman 1990). An example of a betaherpesvirus is human herpesvirus-5 (human cytomegalovirus, HCMV).

Gammaherpesvirinae

The experimental host range of the members of this subfamily is usually limited to the species to which the natural host belongs. *In vitro*, members typically replicate in lymphoblastoid cells and some also cause lytic infection in some types of epithelioid and fibroblastoid cells. Viruses in this group tend to be specific for either B or T lymphocytes, but exceptions occur. In the lymphocyte the infection is frequently either at a pre-lytic or lytic stage, but without production of infectious progeny, and latent virus can be frequently demonstrated in lymphoid tissue. This subfamily was originally subdivided into γ-1-herpesviruses, members of which were typically B-lymphotropic (e.g. human herpesvirus-4, Epstein-Barr virus or EBV) and γ-2-

herpesviruses, members of which were typically isolated from T-lymphocytes (e.g. *Saimirine herpesvirus-2, Herpesvirus saimiri* or HVS), (Honess 1984). Recently the ICTV approved that this subfamily be divided into two genera, namely *Lymphocryptovirus* (formerly γ-1-herpesviruses) and *Rhadinovirus* (formerly γ-2-herpesviruses), (Roizman *et al* 1992).

1.11.5 Biological classification of AHV-1

AHV-1 is classified as a gammaherpesvirus of the rhadinovirus genus. Other viruses in this genus are *Ateline herpesvirus*-2, *Saimirine herpesvirus*-2, *Bovine herpesvirus*-4 and *Murid herpesvirus*-1 (mouse herpesvirus strain, 68) as defined by Roizman *et al* (1992). *Saimirine herpesvirus*-2 and *Ateline herpesvirus*-2 are more commonly referred to as *Herpesvirus saimiri* (HVS) and *Herpesvirus ateles* (HVA) respectively.

The herpesviruses in the rhadinovirus genus do not appear to cause disease in their natural hosts. In susceptible hosts, the histopathological features of AHV-1 infection are very similar to those of known oncogenic viruses such as HVS and HVA, (which might be expected since the these viruses are from the same genus as AHV-1). Biological similarity to EBV (a lymphocryptovirus) was observed (Hunt and Billups 1979), as well as to *Herpesvirus sylvilagus* (Patel and Edington 1980). Patel and Edington (1981) compared AHV-1 to these viruses in terms of the cell-association of virus infectivity, recovery of virus from leucocytes and lymphocytic tissues and the excessive lymphoproliferation associated with such viral infections.

1.11.6 Properties of herpesvirus genomes

The size of herpesvirus DNAs varies from approximately 124 to 235 kilobase pairs (kbp). Independent isolates of the same virus species may also vary in size by as much as 10kbp but this size difference most frequently reflects the number of terminal and/or internal reiterated sequences (Roizman *et al*, 1992).

In base composition the herpesvirus genomes show a broad, but non-exclusive correlation to the biological subgroupings such that *alphaherpesvirinae* have mean compositions of >60% G+C, *betaherpesvirinae* occupy the middle of the range and the rhadinoviruses (formerly γ-2-herpesviruses) account for the majority of viruses with coding sequences of <50% (G+C), (Honess 1984). Herpesvirus DNA also varies with respect to the extent of homogeneity of base sequence distribution along

the genome e.g. human herpesvirus-1 is very homogeneous while HVS has extensive inhomogeneity (Roizman *et al* 1992).

The genomes of gammaherpesviruses are deficient in the dinucleotide CpG, regardless of their G+C contents, as seen in EBV (which has a G+C content of 60%) and HVS (which has a G+C content of 45%), (Baer *et al* 1984, Fleckenstein and Mulder 1980). This depletion probably arises from the mutagenic effects of methylation during latency in a dividing cell (Honess *et al* 1989).

A method of classification of herpesviruses based on their genomic organisations results in six different groupings (A-F), based on the location and nature of the reiterated sequences in the genome, as illustrated in figure 1.1.

AHV-1 has been classified as a member of group B (see figure 1.1) on the basis of its genomic structure (Bridgen *et al* 1989, Roizman *et al* 1992). The arrangement of the DNA in B genomes is such that the terminal sequence is directly repeated numerous times at both termini (Roizman *et al* 1992).

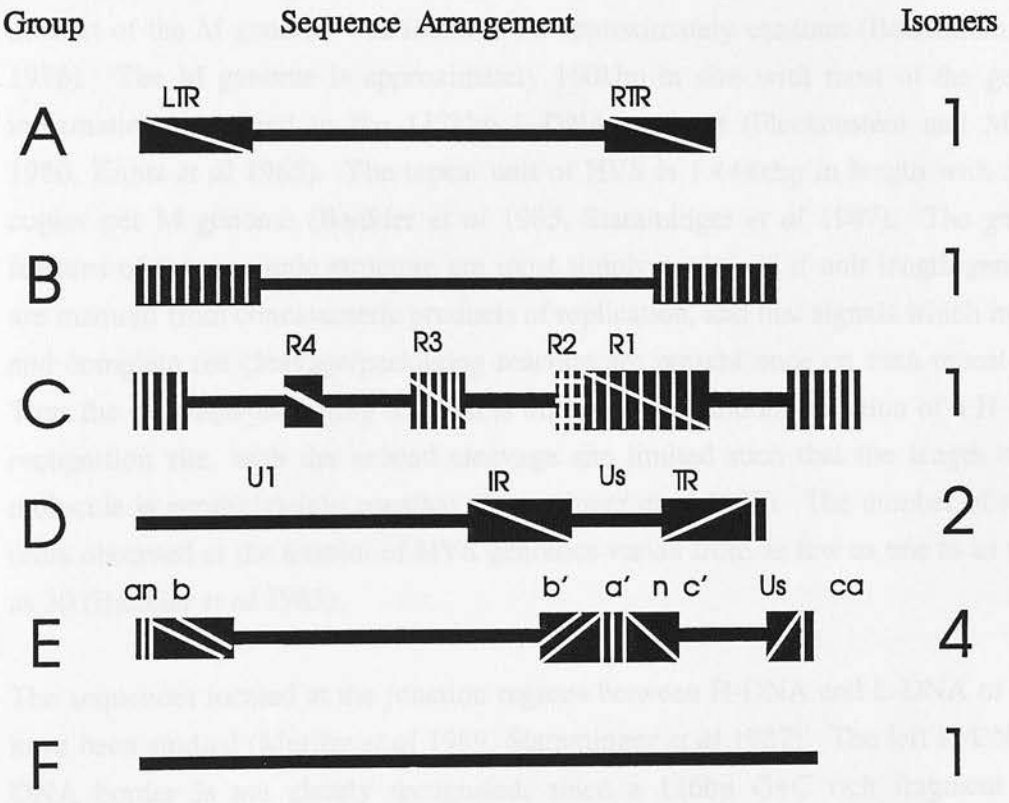
1.11.7 Molecular biology of rhadinoviruses with genomic arrangement B

Herpesvirus saimiri (HVS) is the most extensively characterised member of the rhadinovirus genus. The sequence of the entire HVS genome has been published during the course of this study (Albrecht *et al* 1992), allowing comparison of the AHV-1 and HVS genomes. HVS is regarded as the prototype rhadinovirus (Honess 1984).

HVS is a common infectious agent of squirrel monkeys, which induces malignant lymphomas or acute lymphocytic leukaemias in several new world primate species (Melendez *et al* 1968 and 1969). Caesium chloride density centrifugation of HVS DNA revealed two types of genome. One type, the H genome, consisted of high density reiterated sequences with a G+C content of 71%. The second type, the M genome, consisted of 70% unique sequences referred to as light (L-) DNA, with a G+C content of 36%. The remaining 30% of the M genome was comprised of reiterated sequences of heavy (H-) DNA with a G+C composition of 71%, equal to that of the H genome. The H-DNA of the H genomes and the M genomes was found to be homologous. The H genomes were defective, i.e. no infectious virus was recovered from permissive cells transfected with H-DNA. This was in contrast to the

Figure 1.1

Classification of herpesviruses on the basis of genomic arrangement



Examples

Group	Example
A	ictalurid herpesvirus 1 (channel catfish herpesvirus)
B	saimirine herpesvirus 2 (herpesvirus saimiri)
C	human herpesvirus 4 (Epstein Barr virus)
D	human herpesvirus 3 (Varicella Zoster virus)
E	human herpesvirus 1 (herpes simplex virus)
F	tupaid herpesvirus 1 (tree shrew herpesvirus)

(taken from Roizman *et al* 1992)

cytopathic changes observed in permissive cells transfected with M genomes (Fleckenstein *et al* 1975).

The H-DNA sequences were found to consist of tandem arrays of repeat units which were terminally arranged at both ends of the HVS M genome. The total H-DNA content of the M genome was found to be approximately constant (Bornkamm *et al* 1976). The M genome is approximately 160kbp in size with most of the genetic information contained in the 112kbp L-DNA segment (Fleckenstein and Mulder 1980, Knust *et al* 1983). The repeat unit of HVS is 1.444kbp in length with 30-34 copies per M genome (Bankier *et al* 1985, Stamminger *et al* 1987). The general features of this genomic structure are most simply explained if unit length genomes are matured from concatameric products of replication, and that signals which initiate and complete the cleavage/packaging reaction are present once on each repeat unit. Thus the cleavage/packaging reaction is initiated by a random selection of a H-DNA recognition site, with the second cleavage site limited such that the length of the molecule is approximately constant (Stamminger *et al* 1987). The number of repeat units observed at the termini of HVS genomes varied from as few as one to as many as 30 (Bankier *et al* 1985).

The sequences located at the junction regions between H-DNA and L-DNA of HVS have been studied (Murthy *et al* 1989, Stamminger *et al* 1987). The left H-DNA-L-DNA border is not clearly recognised, since a 116bp G+C rich fragment with homology to repeat DNA lies between concatameric H-DNA repeats and the A+T rich L-DNA sequences (Albrecht *et al* 1992). In contrast the right L-DNA-H-DNA transition was localised to a single nucleotide (Stamminger *et al* 1987). The HVS genome exhibited three small, internal repeat segments in addition to the terminal repeats (Albrecht *et al* 1992).

The other members of the rhadinovirus subfamily generally resemble HVS with regard to their genomic organisation. *Herpesvirus ateles* (HVA) was demonstrated to consist of H and M genomes when analysed by caesium chloride density centrifugation. The H-DNA had a G+C content of 75%, with the M genomes being comprised of L-DNA (37.5% G+C) flanked by approximately 25 identical repeat units. The HVA L-DNA exhibited 35% homology to HVS L-DNA, while the H-DNAs of the two genomes exhibited no homology (Fleckenstein *et al* 1978).

Murid herpesvirus-4 (formerly murine herpesvirus 68) also resembled HVS in that it consisted of a unique region of 118.2kbp flanked by a variable number of copies of a 1.23kbp repeat unit (Efstathiou *et al* 1990b). The G+C content of the unique region of this genome was estimated to be 45% from limited nucleotide sequence analysis (Efstathiou *et al* 1990a).

The *Bovine herpesvirus*-4 (BHV-4) genome is similar to that of HVS and consists of a unique coding region (L-DNA) of 108kbp flanked by (G+C) rich tandem repeats of 1.95, 2.35 and 2.75kbp (H-DNA), (Ehlers *et al* 1985, Bublot *et al* 1990). The genomic organisation of BHV-4 demonstrated overall collinearity with the HVS and EBV genomes (Bublot *et al* 1992). The right junction between the unique and repeated sequences of the BHV-4 genome is relatively constant amongst BHV-4 strains, but the left junction is not so well conserved (Bublot *et al* 1990).

Leporid herpesvirus-1 (*Herpesvirus sylvilagus*), a gammaherpesvirus capable of inducing a lymphoproliferative disease in cottontail rabbits, has also been included as having genomic arrangement B. The genome was found to consist of a unique region of approximately 120kbp, flanked by a variable number of 533bp tandem repeats with a G+C content of 83% (Cohrs and Rouhandeh 1987, Medveczky *et al* 1989a).

Marmodid herpesvirus-1 (woodchuck herpesvirus) has terminally located repeat units of 1.5kbp. The unique region was 130kbp (Gilles and Ogston 1991). This similarity to HVS led to the authors suggesting *Marmodid herpesvirus*-1 should be included in the rhadinovirus genus.

Genomic studies of *Alcelaphine Herpesvirus*-1 have been carried out using the attenuated strain, WC11 (Plowright *et al* 1965). Caesium chloride density centrifugation of this isolate revealed two peaks resembling the M and H DNA components found in HVS (Herring *et al* 1983). The M genome had a G+C content of 50%, and the H genome had a G+C content of 71%. The genome was approximately 160kbp in length, with 130kbp comprising the unique region, and the remaining 30kbp constituting the repeated region. The WC11 genome contained three different sized repeat units. The major repeat unit was 950bp, and was present 15-20 times per molecule, with the two other repeat units, of 1800 and 1050bp respectively, each being present 4-5 times per molecule (Bridgen *et al* 1989). These repeats were terminally located in the WC11 genome (Bridgen 1991).

A second, low passage, cell-associated AHV-1 isolate, C500 (Plowright *et al* 1975) was demonstrated to produce hybridisation profiles similar to those generated for the unique sequence of WC11. The C500 terminal repeats were comprised of multiple copies of a single 1050bp repeat unit (Bridgen *et al* 1989).

1.11.8 Comparison of the genomes of rhadinoviruses and lymphocryptoviruses

Rhadinoviruses and lymphocryptoviruses are both γ-herpesviruses, and the prototypes of each of these genera, HVS and EBV, have been completely sequenced, allowing comparison of the genomic organisation of these two viruses (Albrecht *et al* 1992, Baer *et al* 1984). The two viruses were known to differ in the nature and location of their respective repeated regions, as EBV has the genomic arrangement C in figure 1.1, while HVS follows arrangement B. The genomes differ slightly in size with the EBV genome being 172kbp, and the HVS being 160kbp (Baer *et al* 1984, Albrecht *et al* 1992).

Similarity between the EBV and HVS genomes was summarised by Albrecht *et al* (1992), by demonstrating that the majority of HVS genes have counterparts in the EBV genome at the level of amino acid sequence similarities, and all of these genes are arranged in collinear order. One difference observed between the EBV and HVS genomes involves the genes which encode proteins relevant for transformation and persistance. No HVS equivalents were observed for all the EBV nuclear antigen (EBNA) genes, nor for the latent membrane protein (LMP) or the terminal protein genes (Baer *et al* 1984, Albrecht *et al* 1992). The HVS genome was illustrated to contain a gene which encodes a HVS transforming protein (STP) which was not related to any EBV sequence (Murthy *et al* 1989).

1.12 MOLECULAR CHARACTERISATION OF AHV-1

AHV-1 has been characterised as being distinct from AHV-2 at a molecular level by comparison of restriction endonuclease profiles using six different enzymes (Seal *et al* 1989). Similar comparisons of AHV-1 and OHV-2 have not been possible due to the inability to isolate intact OHV-2. Limited molecular comparison of AHV-1 and OHV-2 was carried out by probing lymphoblastoid cell-line DNA with AHV-1 probes (Bridgen and Reid 1991, discussed previously in section 1.11.1). The OHV-2 fragments recognised varied slightly in size from the AHV-1 fragments used as probes. This illustrated that, while AHV-1 and OHV-2 share homology, the two viruses are distinct.

Shih *et al* (1989) compared the profiles of eight isolates which were associated with MCF using the restriction endonucleases Hind III and Eco RI. These isolates included the laboratory isolates WC11 (which had been recovered from a blue wildebeest) and C500 (which had been recovered from an ox reacting with AHV-1 induced MCF), (Plowright *et al* 1965 and 1975). The other six herpesviruses studied were isolated from captive antelopes. Two distinct patterns were observed using Hind III and Eco RI, exemplified by the WC11 and C500 isolates. The antelope samples resembled either the WC11 or the C500 profiles.

The analysis of AHV-1 DNA as described by Bridgen *et al* (1989) did not identify any six base cutting restriction endonucleases which cleaved the AHV-1 L-DNA from the H-DNA. This was in contrast to the restriction profile observed on Sma I digestion of HVS DNA. Sma I cleaved the H-DNA from L-DNA, but there were no internal Sma I sites in the L-DNA. This facilitated greatly the analysis of the L-DNA with other restriction endonucleases (Knust *et al* 1983).

Limited sequence of the AHV-1 genome was generated by Hsu *et al* (1990b). A 3.5kbp fragment was sequenced, revealing a complete open reading frame (ORF) of 2,058bp and a partial ORF of 630bp. A low degree of homology was observed to EBV. The G+C content of this region of the genome was 47%. Oligonucleotides for use in the polymerase chain reaction (PCR) were designed from this region of the WC11 genome, and as few as 10 infected WC11 cells (representing picogram amounts of AHV-1 DNA) were detected (Hsu *et al* 1990a).

1.13 IDENTIFICATION AND CHARACTERISATION OF THE REGION OF THE HVS GENOME RESPONSIBLE FOR TRANSFORMATION

The HVS rhadinovirus prototype showed strain variability in the leftmost 7-10kbp segment of L-DNA. This region of the genome is highly variable both in size and in nucleotide sequence amongst isolates of HVS (Medveczky *et al* 1984). HVS strain variability led to classification into three groups, A, B and C (originally non-A, non-B) with the rightmost 93% of the L-DNA being almost completely homologous among all strains (Medveczky *et al* 1984 and 1989). HVS strain 11 became attenuated on repeated passage at 39°C, followed by plaque purification at 34°C, generating strain 11att. The 11att genome contained a deletion of 2.3kbp at the left junction between repetitive and unique DNA when compared to the parental strain

(Koomey *et al* 1984). The left terminal variable region was further studied by introducing deletions into a cloned left terminal 7.4kbp fragment, then co-transfecting with the parental strain 11. In the S4 deletion mutants generated, a contiguous fragment of 4.5kbp sequence was found to be unnecessary for *in vitro* culture (Desrosiers *et al* 1984). However, the mutants bearing this deletion were also non-oncogenic. Replacing the sequences lost in both the 11att and S4 mutants also restored the oncogenicity, leading to the conclusion that the left terminal region of the genome was required for oncogenicity (Desrosiers *et al* 1985). The HVS strain SMHI also became attenuated as a result of serial passage in vero cells, generating the variant SMHI-VERO. A region of approximately 6kbp from the 5' terminal region of the genome had been deleted in the non-oncogenic variant. In addition a 5.6kbp fragment had been inserted into the H-DNA. This segment represented an inverted repetition of DNA encoded at the 3' end of the L-DNA.

The complete sequence of HVS strain 11 (subgroup A) illustrates the exact location of the ORF which encodes the protein responsible for the transforming ability of this isolate (Murthy *et al* 1989, Albrecht *et al* 1992). An equivalent ORF in HVS strain 488 (subgroup C) was identified (Biesinger *et al* 1990). This ORF encodes the peptide responsible for the transforming ability of strain 488 (Jung *et al* 1991). The two peptides from strain 11 and strain 488 demonstrated limited structural homology (Jung *et al* 1991). Although this information was not available at the commencement of this study, the work described in the previous paragraph suggested that, if AHV-1 was collinear to HVS, then the region of AHV-1 likely to encode factors responsible for virulent infection of susceptible hosts would be the left terminal L-DNA sequence, when AHV-1 is displayed in the same orientation as HVS.

The map of the AHV-1 WC11 isolate derived by Bridgen (1991) was not complete at the termini of the molecule, nor was it orientated with regard to HVS. In addition the WC11 isolate was known to be attenuated (Plowright *et al* 1965). Comparison of HVS strain 11 and its non-oncogenic variant 11-att revealed a deletion of 2.3kbp from the leftmost region of L-DNA in the attenuated derivative (Koomey *et al* 1984). If attenuation of AHV-1 resulted in a similar deletion then the terminal region of WC11 may lack sequence encoding proteins related to the transforming proteins of HVS. In order to determine the region of the AHV-1 genome responsible for virulence it was necessary to study a virulent isolate, rather than WC11. The C500 isolate of AHV-1 was initially recovered from an ox reacting with MCF (Plowright *et al* 1975). Cell culture propagation of C500 virus results in cell-associated (CA)

virulent virus becoming cell-free (CF) and virulent, then cell-free and attenuated (CFA), (Russell 1980, H.W. Reid, personal communication). Comparison of the virulent and attenuated C500 derivatives at a molecular level may demonstrate variability similar to that observed between the virulent and attenuated variants of HVS.

1.14 PROPOSED OBJECTIVES

The aim of this project was to investigate the genomic rearrangements which result in the loss of virulence of the C500 isolate of AHV-1 during *in vitro* propagation. This was approached initially by examination of restriction profiles of virulent and attenuated C500 derivatives with the intention of identifying any variable regions. Analysis of variable regions may provide evidence of altered reading frames and present an insight into the peptides which are important in establishment of infection and the development of MCF in susceptible hosts.

Chapter 2
Materials and Methods

2.1 SUPPLIERS

All suppliers and their addresses are listed in appendix 1. All chemicals were supplied by either BDH, Fisons or Sigma unless stated otherwise.

2.2 FREQUENTLY USED SOLUTIONS

TE: 10mM Tris, 1mM EDTA; adjusted to pH8.0 with HCl.

20xSSC: 3M NaCl, 0.3M sodium citrate; adjusted to pH7.0 with HCl.

RNase A: 10mg ml^{-1}. Heated to 100°C for 15 min, cooled to room temperature and stored in aliquots at -20°C.

Phenol: Equilibrated with 1M Tris-HCl pH8.0, then with 0.1M Tris-HCl pH8.0. B-hydroxyquinoline added to 0.1% (w/v) and aliquots stored at -20°C.

Chloroform: Chloroform and isoamyl alcohol mixed in a ratio of 24:1 (v/v); stored in a dark bottle at room temperature.

TEG: 25mM Tris-HCl pH8.0, 10mM EDTA, 50mM glucose.

PBS: 137mM NaCl, 26.8mM KCl, 8.1mM Na_2HPO_4, 1.47mM KH_2PO_4.

2.3 GROWTH MEDIA

2.3.1 Bacterial growth media

The quantities of all media components are given in grams per litre final volume of medium, except in cases where final molarities of components are stated.

L-broth: 10g Difco Bacto tryptone, 5g Difco yeast extract, 5g NaCl.

L-agar: as for L-broth, but including 15g Difco agar.

Minimal agar: 0.04M Na_2HPO_4, 0.02M KH_2PO_4, 0.02M NH_4Cl, 0.009M NaCl, 15g agar, 1mM $MgSO_4$, 0.1mM $CaCl_2$, 1mM Thiamine-HCl, 0.2% (w/v) glucose.

Ampicillin: When required media were supplemented with 50µg ml^{-1} ampicillin.

2.3.2 Viral growth media

Virus was grown routinely in Iscove's modification of Dulbecco's medium (IMDM) (Gibco) supplemented with penicillin at 100 IU ml^{-1}, streptomycin at 50µg ml^{-1}, fungazone (Roussel) at 2 IU ml^{-1}, and 5 x 10^{-5}M 2-mercaptoethanol. Foetal bovine serum (FBS) was added at concentrations varying from 5-15% (v/v).

2.4 *Escherichia coli* STRAINS

The *Escherichia coli* strains used in this study are listed in figure 2.1.

2.5 CLONING VECTORS

The cloning vectors used in this study are listed in figure 2.2

2.6 VIROLOGICAL TECHNIQUES

2.6.1 Preparation and storage of viral inoculum

Virulent virus was initially propagated in rabbits by intravenous inoculation of a pool of lymph nodes from a rabbit reacting with MCF following infection with C500. When the animal reacted by demonstrating the classical symptoms of MCF the affected lymph nodes and spleen were harvested. Single-cell suspensions were prepared and stored in IMDM supplemented as described in 2.3.2. These single-cell suspensions were either stored in liquid nitrogen in media containing 15% (v/v) FBS and 0% (v/v) Dimethyl Sulphoxide (DMSO) or inoculated directly onto a bovine cell monolayer and serially passaged. (The animal work and isolation of the virus was carried out by Dr H.W. Reid).

2.6.2 Cultivation of virus

Viral isolates were propagated in monolayer cultures of bovine thyroids and bovine turbinates, both of which are low passage cells . All isolates were cultivated at 37°C in IMDM as described in 2.3.2, with 5% (v/v) FBS added. Cultures were harvested when approximately 50-80% of the monolayer exhibited cytopathic effects (cpe).

2.6.3 Assessment of virulence

Bovine turbinate monolayer cultures in 25cm^2 flasks were infected with the test virus and cultured at 37°C until approximately 50% of the monolayer exhibited cpe. The cells were scraped from the plastic and resuspended in approximately 2ml of the supernatant fluid. One ml was removed for inoculation into a rabbit while the remainder was mixed with an equal volume containing 20% DMSO and 50% FBS prior to storage in liquid nitrogen.

Rabbits were inoculated into the lateral ear vein with the test virus and examined daily for clinical signs and rectal temperature. Rabbits were considered to have

Figure 2.1

Escherichia coli strains

Strain	Genotype	Reference
JM101	*sup*E *thi*-1 Δ(*lac-pro*AB) F' [*tra*D36 proAB$^+$ *laq*I^q *lac*ZΔM15]	Yannisch-Perron *et al* 1985
JM109	*rec*A1 *sup*E44 *end*A1 *hsd*R17 *gyr*A96 *rel*A1 *thi* D(*lac-pro*AB) F' [*tra*D36 proAB$^+$ *laq*I^q *lac*ZΔM15]	Yannisch-Perron *et al* 1985

Figure 2.2

Vector	Type	Size	Replicon	Markers	Promoters
pBS+	RNA expression ss DNA	3200bp	ColE1	amp, *lacZ'*	T3, T7
pBS SK+	RNA expression ssDNA	2950bp	ColE1, f1	amp, *lacZ'*	T3, T7

Vector	Multiple cloning sites
pBS+	Hind III, Sph I, Pst I, Hind II, Acc I, Sal I, Xba I, Bam HI, Xma I, Sma I, Kpn I, Sac I, Eco RI
pBS SK+	Sac I, Bst I, Sac II, Not I, Eag I, Xba I, Spe I, Bam HI, Sma I, Pst I, Eco RI, Eco RV, Hind III Hind II, Acc I, Sal I, Xho I, Apa I, Dra II, Kpn I

References:
pBS+ Yannisch-Perron *et al* (1985)
pBS SK+ Short *et al* (1988)

reacted when the rectal temperature exceeded 40°C, at which time there was also a marked reduction in the quantity of food and water ingested. On the second day of reactin rabbits were killed by gaseous anaesthesia and examined for gross pathological changes. Rabbits which did not react by the 90th day were considered not to have beeen infected and the experiment was discontinued.

2.6.4 Preparation of DNA from cell-associated (CA) virus

CA virus was harvested when over 80% of the cells exhibited cpe. Since the cpe resulted in the majority of the monolayer cells becoming rounded or forming loosely attached syncytia these cells were easily incorporated into the supernatant and any residual cells were removed from the flask walls via rocking briefly with sterile glass beads. The culture supernatant was then spun at 3000g for 10 min at room temperature (RT) to pellet the cells, which were then washed in PBS. The cell pellets were resuspended and lysed in 0.5% (v/v) Nonidet P40 (NP40) in RSB (10mM NaCl, 1mM $MgCl_2$, 10mM Tris-HCl, pH7.5) by gentle homogenisation. Chromatin and cellular debris were removed by centrifugation at 1000g for 10 min at RT. The supernatant was then layered over 25% sucrose in RSB and the viral nucleocapsids were isolated by centrifugation at 114,000g for 30 min at RT. The resultant pellet was resuspended in TNE (10mM Tris-HCl, 100mM NaCl, 1mM EDTA, pH8.0) and digested with proteinase K (Boehringer Mannheim) at 200µg ml^{-1} in the presence of 1% SDS for 2 hr at 37°C with agitation. Contaminants were extracted with phenol and the DNA precipated with ethanol.

2.6.5 Preparation of DNA from cell-free (CF) virus.

CF virus was harvested when over 50% of the cells exhibited cpe. The cells were detached with glass beads and the culture supernatant was first clarified by centrifugation at 4,000g for 20 min at 4°C, then pelleted by centrifugation at 42,000g for 30min at 4°C. The pellet was resuspended in TNE and thereafter treated with proteinase K (Boehringer Mannheim) and extracted following the method described for CA virus.

2.7 BACTERIOLOGICAL TECHNIQUES

2.7.1 Growth and maintenance of E. coli

Bacteria were grown and maintained on solid and liquid media using standard techniques as described by Maniatis *et al* (1982). Cells were maintained on L-agar or minimal agar plates for 2-3 weeks at 4°C, while longer term storage was achieved by

freezing at -70°C in L-broth supplemented with a final concentration of 15% glycerol.

2.7.2 Preparation of competent cells

Competent bacterial cells were prepared by either of two methods:

i) Fresh competent cells were prepared according to an adaptation of the method described in Maniatis *et al* (1982), based on the original method of Mandel and Higa (1970).

An overnight culture of one of the *E. coli* strains described in figure 2.1 was diluted 1:50 in L-broth and grown at 37°C with agitation until an OD_{600} of 0.5 was reached. The cells were cooled on ice for 10 minutes, harvested by centrifugation at 3000g for 10 min and resuspended in 0.5 volumes of cold (4°C) sterile 75mM $CaCl_2$. The cells were incubated on ice for 30 min then pelleted as before and resuspended in 0.08 volumes of cold sterile 75mM $CaCl_2$. The cells were left on ice for at least an hour before use and were kept for a maximum of 5 days at 4°C or for 2-3 months at -70°C, after the addition of glycerol to a final concentration of 10%.

ii) Competent cells were also prepared and stored by the method of Chung *et al* (1989).

An overnight culture was again diluted 1:50 in L-broth and grown to an OD_{600} of 0.3-0.4. The cells were cooled on ice for 15 min then harvested by centrifugation at 3000g for 10 min. The pellet was resuspended in 0.1 volumes of L-broth containing 10% (w/v) PEG (molecular weight approximately 8000), 5% DMSO and 25mM $MgCl_2$. Cells were then stored in aliquots at -70°C for 4-6 months.

2.8 NUCLEIC ACID AND MOLECULAR TECHNIQUES

2.8.1 Small scale preparation of plasmid DNA

Double stranded plasmid DNA was prepared from 10ml overnight cultures grown under antibiotic selection by the method described by Maniatis *et al* (1982), based on the original method of Birnboim and Doly (1979).

1.5ml of the culture was pelleted by centrifugation at 9000g for 30 sec then resuspended in 100µl TEG. 200µl of a freshly prepared solution of 1% (w/v) SDS, 0.2M NaOH was then added to the contents of the tube, mixed by inversion then incubated for 5 min on ice. Cellular debris, protein and genomic DNA were then

precipitated by the addition of 150µl of 3M Potassium and 5M acetate, pH4.8, for 5 min on ice then pelleted by centrifugation at 9000g for 5 min. The supernatant containing the supercoiled plasmid DNA was extracted with phenol:chloroform (6:4) then precipitated by the addition of 1ml of ethanol. The DNA was resuspended in 50µl of TE.

2.8.2 Large scale preparation of plasmid DNA

A single bacterial colony was inoculated into 5µl of L-broth containing 50mg ml^{-1} ampicillin and grown overnight at 37°C with shaking. This culture was then used to inoculate 500ml of L-broth (containing 50µg ml^{-1} ampicillin) and grown overnight at 37°C with shaking. Bacterial cells were harvested by centrifugation at 4000g for 10 min at 4°C. Cells from 250ml of culture were lysed by resuspension in 20ml TEG containing 1mg ml^{-1} of freshly prepared lysozyme and incubated for 10 min on ice, followed by the addition of 40ml of freshly prepared 0.2M NaOH, 1% (w/v) SDS and a further 10 min incubation on ice. Chromosomal DNA, high molecular weight RNA and protein/membrane/SDS complexes were precipitated by the addition of 3M potassium and 5M acetate, pH4.8. The precipitate was pelleted by centrifugation at 22,000g for 10 min at 4°C, and the supernatant recovered. Plasmid DNA was recovered by the addition of 0.6 volumes of propan-2-ol and precipitated by centrifugation at 15,000g for 15 min at RT. The pellet was subsequently washed in 70% ethanol and resuspended in 8ml TE. Exactly 1g of caesium chloride was added to every ml of this solution and 0.8ml of ethidium bromide solution (10mg ml^{-1}) was then added for every 10ml of the caesium chloride-DNA solution. This resulted in a final density of 1.55g ml^{-1}. Protein and cellular debris were removed from this solution by centrifugation at 3000g at RT. The supernatant was then transferred into Beckman "quickseal" tubes and centrifuged to equilibrium at 165,000g in a Beckman Ti70 rotor for 48 hr at 20°C. After centrifugation two bands located near the centre of the tube were visible in ordinary light. The upper band, consisting of chromosomal and nicked plasmid DNA, was collected first to prevent contamination of the lower band, using an 18-gauge hypodermic needle. The lower band, containing closed circular plasmid DNA was then collected in the same way. Ethidium bromide was removed from the plasmid DNA solution by extracting with water-saturated butan-1-ol, then the DNA was precipitated by the addition of 1 volume of water and 6 volumes of ethanol.

2.8.3 Quantification of DNA

The concentration of DNA preparations was assessed by a measurement of OD_{260} using a spectrophotometer. An OD_{260} of 1 corresponds to 50µg ml^{-1} double stranded

DNA. DNA preparations free from protein possessed an OD_{260}:OD_{280} ratio approaching 2.0.

2.8.4 Restriction endonuclease digestion of DNA

DNA was digested with restriction endonucleases at a final dilution of approximately 10 units μg^{-1}. All restriction endonucleases and their relevant 10X buffers were purchased from Boehringer Mannheim and used as per manufacturers instructions. RNase was added at a final concentration of 20μg ml^{-1} when required and the volume made up to between 10 and 20μl with sterile distilled water. The reactions were incubated for 90 min at the temperature recommended by the manufacturer.

2.8.5 Ligation of DNA

Ligation reactions were prepared using appropriate amounts of vector DNA, insert DNA (at a ratio of 1:4 vector:insert where possible), 10X concentrated ligation buffer (660mM Tris-HCl, 50mM $MgCl_2$, 10mM dithiothrietol, 10mM ATP, pH7.5) and 1-3 units of T4 ligase (Boerhinger Mannheim). Reactions were carried out in a volume of 5-20μl (made up with sterile distilled water) and incubated for 1-4 hr at RT or for 16-20 hr at 12°C.

2.8.6 Dephosphorylation of DNA

The terminal 5' phosphates were removed from DNA by treatment with calf intestinal alkaline phosphatase (CIP), (Boehringer Mannheim). Dephosphorylation reactions were prepared using target DNA, 10X CIP buffer (0.5M Tris-HCl pH9.0, 10mM $MgCl_2$, 1mM$ZnCl_2$, 10mM spermidine) and 0.01 units of CIP per μg of DNA in a final volume of 10-20μl (made up with sterile distilled water). Incubations were carried out according to Sambrook *et al* (1989) with protruding 5' termini being incubated for 30 min at 37°C, while blunt or recessed termini were first incubated for 15 min at 37°C, after which a second aliquot of CIP was added and the incubation carried out for a further 45 min at 55°C. Following incubation the CIP was inactivated by heating for 10 min at 75°C in the presence of 5mM EDTA prior to extraction with phenol:chloroform then chloroform. The DNA was precipated by adding 0.1 volumes 3M Sodium Acetate and 2 volumes ethanol.

2.8.7 Transformation of competent E. coli

Up to 50ng of ligated plasmid DNA was added to 200μl of competent cells. The DNA and cells were subsequently incubated for 30 min on ice, then heat-shocked for 4 min at 37°C. 800μl of L-broth was then added and the cells incubated for 1 hr at 37°C to allow expression of antibiotic resistance. The cells were pelleted and

resuspended in 200µl of L-broth with aliquots of 30 and 80µl then spread on L-agar plates containing 50µg ml^{-1} ampicillin, 24µg ml^{-1} isopropyl -D-thiogalactopyranosidase (IPTG),(Nova Biochem), and 20µg ml^{-1} X-galactose (X-gal), (Nova Biochem), for selection of recombinants.

2.8.8 Transfer of bacterial colony DNA to nitrocellulose

Bacterial colonies were transferred to nitrocellulose and the DNA bound by the method described in Maniatis *et al* (1982) with modifications.

Colonies selected as containing recombinant plasmids using IPTG and X-gal (i.e. white colonies as opposed to blue) were screened by transferring to a fresh L-agar plate and a replica L-agar plate, overlaid with a nitrocellulose filter, using sterile toothpicks. The filters used were Hybond-C extra (Amersham). Both plates were incubated at 37°C overnight. The filter was then treated by placing over 3MM Whatman chromatography paper impregnated with the following solutions. First the filter was treated with 0.5M NaOH, 1.5M NaCl for 7 min, then twice with 1M Tris-HCl (pH7.4) for 2 min, and finally with 1M ammonium acetate for 4 min. Once dried the filter was baked in a vacuum oven at 80°C for 1 hr, then washed in 5X SSC, 0.5% (w/v) SDS, 1mM EDTA for 30 min at 50°C. Cellular debris was removed by rubbing the filters gently with dampened tissues.

2.8.9 Agarose gel electrophoresis of DNA

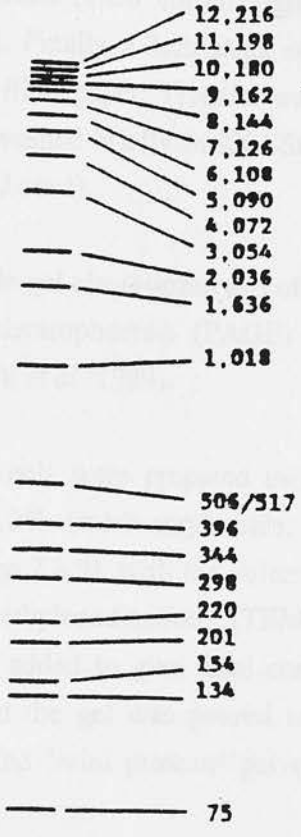
Nucleic acid grade "ultrapure" agarose (BRL) was used at concentrations of 0.6-1.2% (w/v) in either Loening "E" buffer (180mM Tris, 150mM NaH_2PO_4, 5mM EDTA, pH7.6-7.8) or TAE buffer (40mM Tris-acetate, 2mM EDTA pH7.5-7.8). Ethidium bromide was added to molten agarose at a final concentration of 0.2mg ml^{-1}. 0.1 volumes of DNA loading buffer (0.1% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 75mM EDTA, 25% sucrose) was added to DNA samples prior to loading, and electrophoresis was carried out at 1-4 volts cm^{-1} in the relevant buffer until the required resolution was achieved. Size estimation was achieved by co-electrophoresis of a 1kbp ladder (GIBCO-BRL), and the fragment sizes of this ladder are illustrated in figure 2.3. Agarose gel tanks were purchased from Biorad and Pharmacia. DNA fragments were visualised by fluorescence of bound ethidium bromide in ultra violet (uv) light of 302nm wavelength.

2.8.10 Southern transfer of DNA to nylon filters

Transfer of DNA from agarose gels to nylon filters was carried out following a modification of the method of Smith and Summers (1980).

Figure 2.3

Fragment sizes of the 1kbp ladder



The gel was incubated in 0.25M HCl for 10 min to depurinate the DNA (if fragments did not exceed 15kbp the depurination step was omitted) then rinsed in distilled water and incubated in denaturing solution (0.5M NaOH, 1.5M NaCl) for 30 min. The gel was subsequently neutralised by incubation in 1M Ammonium acetate for 1 hr. Transfer to "Hybond N" nylon membrane (Amersham) was carried out by placing the membrane (previously soaked in neutralising solution) on top of the gel, followed by 3 sheets of Whatman 3MM chromatography paper (also previously soaked in neutralising solution). Finally a 3cm stack of paper towels and a 1kg weight were placed on top of the filter paper. Transfer was allowed to proceed for 3-18 hr after which the filter was washed briefly in 2X SSC, air dried and fixed to the membrane by uv irradiation (0.4 J cm^{-1}).

2.8.11 Polyacrylamide gel electrophoresis of DNA

Polyacrylamide gel electrophoresis (PAGE) of DNA was based on the method described by Sambrook *et al* (1989).

7.5% polyacrylamide gels were prepared using appropriate volumes from a 30% acrylamide stock (29.2% (w/v) acrylamide, 0.8% (w/v) bisacrylamide) and 5X Loening "E" buffer (see 2.8.9), with the volume being made up with distilled water. N,N,N',N',-tetramethylethylenediamine (TEMED) and ammonium persulphate (APS), (Biorad) were added to give final concentrations of 0.1% (v/v) and 0.1% (w/v) respectively, and the gel was poured immediately after their addition. Gels were cast and run in the "mini protean" gel system (Biorad) at 200 volts constant voltage setting.

2.8.12 Silver staining of polyacrylamide gels

Nucleic acid fragments were silver stained using a modification of the technique described by Herring *et al* (1982).

Following electrophoresis the gel was fixed in 100ml of 10% (v/v) ethanol and 0.5% (v/v) ethanoic acid. The gel was then stained in 11.2mM $AgNO_3$ for 10 min. After a quick wash in distilled water the bands were developed by the addition of a solution of 0.75M NaOH containing 0.25% (v/v) formaldehyde for 2-5 min. When the desired intensity was achieved the developing solution was removed and the reaction stopped by washing in 75mM Na_2CO_3. All the steps in this staining procedure were carried out with gentle agitation.

2.8.13 Recovery of DNA from agarose gels

DNA fragments were recovered from agarose gels by either of two methods depending on the size of the fragment:

i) DNA fragments were recovered from agarose gels using powdered glass following the technique described by Vogelstein and Gillespie (1979). This process was carried out using the commercially available "Gene clean 2" kit (Stratech Scientific Ltd).

Briefly, an agarose slice containing the required DNA was excised from the gel and dissolved in 2-3 volumes of 3M NaI at 50-55°C. 5µl "glass-milk" (a specially formulated silica matrix in water) was then added and the contents of the tube were vortexed then left for 5 min at RT to allow the DNA to bind to the glass. The DNA-glass matrix was then pelleted at 9000g for 5 sec and washed in "NEW wash" (a solution containing NaCl, Tris, EDTA and ethanol) 3 times. The DNA was eluted in 5µl of sterile distilled water at 50°C for 3 min and the "glass-milk" pelleted by centrifugation at 9000g for 30 sec. This technique was used for fragments greater than 300bp.

ii) DNA fragments were recovered from agarose gels using the commercially available agarase (GELase, Cambio Ltd).

GELase was used to recover DNA fragments from low melting point (LMP) agarose (BRL) gels of 1.5-2.0% agarose. The gel slice containing the required DNA was excised from the gel and 50X GELase buffer was added at a concentration of 2µl per 100µg gel slice. After incubation at 65-70°C to melt the agarose GELase was added according to the manufacturers instructions, and the tube was incubated at 40°C. The DNA was recovered by the addition of 1 vol of 5M ammonium acetate and 2 volumes of ethanol followed by centrifugation at 9000g for 10 min.

2.8.14 Preparation of digoxygenin labelled DNA probe

Probe DNA was labelled with digoxygenin (DIG) using the commercially available "DIG DNA labelling and detection kit" (Boehringer Mannheim).

DNA was prepared following the methods described in 2.8.4, 2.8.9 and 2.8.13 where appropriate, then was labelled by random primed incorporation (Feinberg and Vogelstein 1983) of DIG. The DNA was denatured by boiling for 10 min, then incubated with 1µl random hexanucleotide mix, 1µl dNTP labelling mix and 1 unit

of klenow enzyme for 1-24 hr at 37°C. This labelling reaction results in incorporation of DIG-11-dUTP every 20-25 nucleotides in the newly synthesised DNA.

2.8.15 Preparation of digoxygenin labelled DNA probe using specific primers

An adaptation of the manufacturers recommended labelling reaction for the "DIG DNA labelling and detection kit" was used to label small DNA fragments (50-500bp) cloned into plasmids (either pBS+ or pBS SK+) containing M13 universal and reverse primer sequences.

The target plasmid was first linearised by restriction to prevent labelling of plasmid DNA beyond the insert. The DIG labelling reaction was carried out by adding 1µl of the relevant sequencing primer, 1µl of dNTP mix and 1 unit of klenow enzyme to the linearised plasmid in a final volume of 10µl. The reaction was incubated at 37°C for 1-24hr. This adaptation was used to generate more specific probe (since priming is at one site only) and to avoid losses of small fragments incurred during recovery from agarose.

2.8.16 Preparation of digoxygenin labelled RNA probe

Probe RNA was labelled using digoxygenin-11-uridine-5-triphosphate (DIG-11-UTP) (Boehringer Mannheim).

This method was used to generate probes of 100-600bp from fragments cloned into pBS+ or pBS SK+, both of which contained recognition sequences for the T3 and T7 RNA promoters. The template DNA was first linearised and denatured. Labelling with DIG-11-UTP was carried out by adding 3.5mM DIG-11-UTP to 6.5mM UTP, 10mM ATP, 10 mM GTP and 10mM CTP to produce a 10X "NTP mix". Forty units of T3 or T7 RNA polymerase (Boehringer Mannheim) and 2µl of 10X transcription buffer (0.4M Tris-HCl, pH8.0, 60mM $MgCl_2$, 100mM dithiothreitol, 20mM spermidine, 100mM NaCl) were added to the linearised DNA, along with 2µl 10X "NTP mix", 2µl RNAse inhibitor (Boehringer Mannheim) and sterile water to make the volume up to 20µl. The reaction was incubated for 2 h at 37°C and the probe then frozen at -20°C until required. This method of labelling was used to give higher yields of probe.

2.8.17 Hybridisation of DIG labelled probes

Hybridisations were carried out essentially as recommended by the manufacturers (Boehringer Mannheim).

Filters were prehybridised by incubating with hyridisation solution (5X SSC, 5% (w/v) "blocking reagent", 0.1% (w/v) N-lauroylsarcosine, Na-salt, 0.02% (w/v) SDS, 50% (v/v) formamide). This solution was then removed and replaced with 2.5ml per $100cm^2$ of fresh hybridisation solution containing freshly denatured probe. Hybridisation was carried out for 16-20 hr at 42°C. Filters were washed twice in 2X SSC, 0.1% (w/v) SDS for 5 min at RT, then twice in 0.1X SSC, 0.1% (w/v) SDS for 15 min at 68°C.

2.8.18 Detection of bound DIG labelled probes

Immunological detection of bound DIG labelled probe was carried out according to the manufacturers instructions with minor alterations. All steps were carried out with gentle agitation unless otherwise stated.

Filters were blocked in a solution of TN buffer (100mM Tris-HCl, 150mM NaCl) containing 0.5% (w/v) "blocking reagent" for 30 min. The filters were then incubated in 10-20ml TN buffer containing a 1:5000 (v/v) dilution of anti-DIG alkaline-phosphatase conjugate. Unbound antibody-conjugate was removed by washing twice in TN buffer for 15 min. The filters were then equilibrated in TNM (100mM Tris-HCl, 100mM NaCl, 50mM $MgCl_2$, pH9.5) before adding the colour solution which consisted of 45µl "NTB" solution and 35µl "X-phopsphate" in 20ml TNM. ("NTB" consists of 75mg ml^{-1} nitroblue tetrazolium salt in dimethylformamide while "X-phosphate" consists of 50mg ml^{-1} in 5-bromo-4-chloro-3-indolyl phosphate, toludinium salt, in dimethylformamide). The filters were left static in the dark for 1hr to 3 days until the colour had developed to the required intensity. The reaction was stopped by washing briefly in TE, after which the filters were air dried.

2.8.19 Di-deoxy chain termination sequencing of DNA using T7 polymerase

T7 polymerase sequencing was carried out using the commercially available T7 sequencing kit (Pharmacia), which provided all the reagents necessary to carry out DNA sequencing reactions by the di-deoxy termination method of Sanger *et al* (1977), except for [^{35}S] dATP, which was obtained from Amersham. Preparation of DNA was as described in 2.7.1-2, and all sequencing was carried out using double stranded DNA as template. All procedures were carried out as recommended by the manufacturers.

Approximately 1µg per reaction was first denatured by adding 0.4M NaOH and incubating for 10 min at RT. The DNA was then precipitated by the addition of 0.3 volumes of 3M sodium acetate 0.7 volumes distilled water and 6 volumes of ethanol.

2µl of "annealing buffer" (a buffered solution containing $MgCl_2$ and dithiothreitol) and 2µl of primer (1.6mM) were added to the DNA, mixed gently, then incubated for 20 min at 37°C to anneal primer to template, or heated to 65°C for 10 min then left to anneal while cooling to RT . Once the annealing reactions had cooled the labelling and termination reactions were carried out. An "enzyme premix" was prepared by adding 3 units of T7 polymerase in "enzyme dilution buffer (a buffered solution containing glycerol, bovine serum albumin and dithiothreitol) to 1µl (10µCi) of [^{35}S] dATP, 3µl of "labelling mix A" (dCTP, dGTP and dTTP in solution) and 1µl of distilled water. 6µl of "enzyme premix" was added to the annealed template and primer and incubated for 5 min at RT during which time the newly synthesised DNA was labelled by incorporation of [^{35}S] dATP. Chain termination was effected by the addition of 4µl from this reaction to each of 4 tubes containing 2.5µl of "G", "A", "T", and "C" mixes respectively (each mix contained deoxy and di-deoxy forms of the respective base) and incubating for 5 min at 37°C. 5µl of "stop solution" was added to each tube prior to storage at -20°C.

The sequencing samples were resolved by PAGE. 6% acrylamide gel solution was prepared from a 40% acrylamide stock (38% (w/v) acrylamide, 2% (w/v) bis-acrylamide) with urea added to a final concentration of 7M and 10X TBE (1M Tris-HCl, 0.865M boric acid, 20mM EDTA) at a dilution of 1:10. 0.1% (w/v) APS and 0.05% (v/v) TEMED were added immediately before pouring the gel. Electrophoresis in 1X TBE buffer was carried out using S2 (BRL) apparatus at 50 watts constant power setting. The gels were pre-run for 30 min before loading freshly denatured samples (heated to 80°C for 2 min). Further samples were added as required and electrophoresis continued for 5-8 hr. The gels were fixed by immersing in 10% (v/v) methanol, 10% (v/v) ethanoic acid then transferred to Whatman 3MM filter paper and dried under vacuum at 76°C for 30 min.

Detection of [^{35}S]-labelled nucleic acids in sequencing gels was achieved by exposure to medical x-ray film (Fuji) in radiography cassettes. The x-ray film was developed and fixed in "Photosol" CDL8 developer and "Photosol" CF40 fixer (Photosol products were purchased from Genetic Research Instrumentation Limited).

2.8.20 Polymerase chain reaction

Amplification reactions were performed according to the procedure described by Saiki *et al* (1985 and 1988), following guidelines described by Innis *et al* (1990) with minor modifications.

Polymerase chain reactions (PCR) were set up by adding final concentrations of 0.2mM dNTPs, 0.25mM of two 20bp oligonucleotide primers, 1X PCR buffer (10mM Tris-HCl, 1.5mM $MgCl_2$, 50mM KCl, 0.1mg ml^{-1} gelatine) and 2.5 units Taq polymerase (Boehringer Mannheim) to freshly denatured target DNA in a final volume of 50µl. The target DNA was 100ng of PP1, PP2, CA or CFA genomic DNA. The reactions were overlaid with 50µl of sterile mineral oil before being amplified using a Perkin Elmer Cetus Thermal DNA Cycler or a Hybaid ampligene temperature cycler.

The PCR conditions varied according to several factors. The annealing temperature was determined by calculating the T_m of both oligonucleotides, with each A and T accounting for 2°C, and each G and C accounting for 4°C. Annealing was carried out at 5°C below the lower T_m of the two oligonucleotides. The extension phase of the cycle varied depending on the predicted size of the product, with 1 minute being added on for every 1kbp synthesised. All PCR reactions followed a cycle of denaturation at 94°C for 1 minute, annealing at various temperatures for 1 minute, and extension at 72°C for varying times. Figure 2.4 lists the variable factors involved in these reactions.

PCR fragments were analysed by agarose gel electrophoresis or polyacrylamide gel electrophoresis depending on the size of the fragment. Confirmation that the correct fragment was produced on PCR cycling was achieved by Southern transfer and subsequent hybridisation using DIG labelled probes.

2.8.21 The use of PCR to generate DIG-labelled probes

PCR was used to prepare highly specific DIG-labelled probes in large quantities. The PCR reactions were set up and carried out as described in 2.8.20, with DIG-dUTP incorporated into the 2mM dNTP 10X stock at a ratio of 35:65 (DIG-dUTP:dTTP). The reaction then proceeded following a typical cycle as described in 2.8.20, repeated fifteen times. The resultant probe was analysed by agarose electrophoresis to assess the purity and either recovered from agarose (following either of the methods described in 2.8.13) or used directly.

2.8.22 Di-deoxy chain termination sequencing of DNA using Taq polymerase

Taq polymerase sequencing of DNA was carried out using the commercially available fmol DNA sequencing system (Promega). This method employs PCR to prepare samples using the di-deoxy termination method of Sanger *et al* (1977). The sequencing reactions carried out all involved incorporation of [^{35}S] dATP

Figure 2.4

Reaction	Annealing Temperature	Extension time (+ extension per cycle)	Number of cycles
A	50°C	3min (+ 10sec)	25
B	45°C	3min (+ 10sec)	30
C	45°C	3min (+ 30sec)	30
D	55°C	3min (+ 10sec)	30
E	45°C	1min 30sec	30
F	45°C	2min (+ 5sec)	30
G	40°C	1min	15
H	55°C	1min	30

(Amersham) and followed the manufacturers instructions, with all additional reagents supplied in the kit.

Approximately 1µg of plasmid DNA was used per reaction, and to this was added 3.0pmol primer, 10µCi [^{35}S] dATP, 5µl 5X fmol sequencing buffer (250mM Tris-HCl, 10mM $MgCl_2$), 5U sequencing grade Taq polymerase and sterile dH_2O to a final volume of 17µl. 4µl of this was added to each of 4 tubes containing 2µl of 'G', 'A', 'T' and 'C' nucleotide mixes. The nucleotide mixes were comprised of deoxy and di-deoxy nucleotides, with the dGTP being 7 deaza dGTP to reduce hydrogen bonding and therefore resolve compressions. The sequencing reactions were subsequently carried out following a cycle of 95°C for 30 seconds, 45°C for 30 seconds, then 72°C for 1 minute. After 30 cycles 3µl of "stop solution" (10mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanole) was added to each tube. The sequencing reactions were resolved by PAGE in an identical manner to that described in 2.8.19.

Chapter 3
Identification, cloning and characterisation of variable regions of C500

3.1 INTRODUCTION

The C500 isolate of AHV-1 was derived from the blood of an ox reacting with MCF following presumptive exposure to wildebeest (Plowright *et al* 1975). This isolate of AHV-1 was shown to retain virulence for a limited number of passages in tissue culture, in both cell-associated and cell-free forms (Russell 1980). The C500 isolate became attenuated (i.e. failed to induce typical MCF symptoms following inoculation into rabbits) on continual serial passage in tissue culture (unpublished data). This allowed comparison of virulent and attenuated derivatives of a single AHV-1 isolate at a molecular level. This chapter describes the derivatives of C500 employed in this study and the initial experiments carried out to compare these derivatives.

3.2 C500 TISSUE CULTURE DERIVATIVES

The C500 virus was typically cell-associated (CA) and virulent for the first fifteen to twenty passages in monolayered cells of bovine lineage. Virus infectivity subsequently was largely present in a cell-free (CF) form but retained virulence for a relatively transient phase of approximately five passages before losing pathogenicity for rabbits. The cytopathic effects (cpe) of both the virulent and the attenuated C500 derivatives on bovine turbinate cells are illustrated in figures 3.1a and 3.1b respectively. Infection with the virulent, cell-associated virus typically results in the formation of syncytia and a few rounded cells, whilst infection with the attenuated, cell-free virus results in single rounded cells with few or no syncytia. Infection with both C500 derivatives finally resulted in deterioration of the bovine monolayer. With cell-associated virus this generally occurred 4-6 days after infection, compared to 2-4 days with cell-free virus. The different cpe observed on infection with these two C500 derivatives is consistent, implying that the virus has undergone biological alterations.

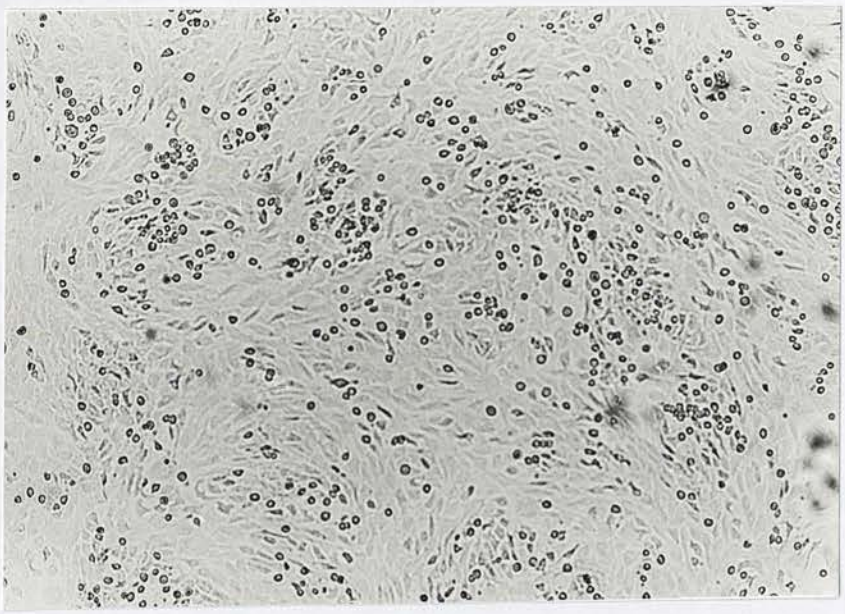
The study of the genomic alterations which take place in the switch from virulence to attenuation was performed by comparison of the C500 derivatives. When this study began, the C500 virus was available as virulent, CA virus, virulent CF virus and attenuated CF virus (CFA). In addition to these three derivatives low passage C500 CA virus was biologically cloned by three serial limiting dilutions. This population was subsequently denoted PP. Comparative studies of AHV-1 also included (where appropriate) WC11, a strain which was completely

Figure 3.1

Infection of bovine turbinate cells with C500 virus

Figure 3.1a Infection with virulent, cell-associated C500 virus (magnification X100).

Figure 3.1b Infection with attenuated, cell-free C500 virus (magnification X60).



Both cultures were photographed 48hr post infection

attenuated following prolongued *in vitro* propagation (Plowright *et al* 1965). The five AHV-1 derivatives compared in this study are summarised in figure 3.2.

3.3 MOLECULAR COMPARISON OF THE C500 DERIVATIVES

Purified DNA from the four C500 derivatives was compared by restriction endonuclease (RE) digestion using five enzymes which had six-base recognition sites. To provide a high degree of resolution and to allow small differences to be observed the digests were resolved using polyacrylamide gel electrophoresis and visualised by silver staining.

The RE used to compare the four C500 derivatives in the first instance were Hind III, Eco RI, Bam HI and Xho I. Digestion with these enzymes failed to identify any appreciable differences in the restriction profiles of the four C500 derivatives (data not shown). This result was predictable since the aforementioned enzymes all lack sites in the terminal repeat region. The termini of the C500 molecule, as described in section 1.11.7, consist of a variable number of repeats of a 1050bp unit. In a given population of virions the number of repeat units at each end of the genome can vary significantly, hence an RE which lacks sites in the repeat unit produces terminal fragments of variable sizes. The result of using such enzymes is that a ladder of fragments is formed for each end of the molecule. Hind III, Eco RI, Bam HI and Xho I all must lack sites in the repeat unit as all produced double ladders which caused a masking effect thus making the profiles of these enzymes difficult to interpret.

When the four C500 derivatives were digested with Sma I, an enzyme which recognises sites in the repeat unit, two differences were observed on silver staining, as illustrated in figure 3.3. Firstly, a fragment of approximately 5kbp, present in the three virulent derivatives i.e. PP, CA and CF, was absent from the CFA derivative. Secondly, a fragment of approximately 3.8kbp was observed in the CA, CF and CFA derivatives but not in the PP derivative. The observation that the 5kbp fragment was present in all the virulent derivatives but was absent from the attenuated derivative suggested that loss of this fragment may be related to loss of virulence. Before this possibility could be further examined it was necessary to ascertain whether loss of this fragment was a genuine feature of attenuation or an artefact of one particular CFA derivative. A second, independent C500 attenuated derivative, CFA 2, was digested with Sma I, as was WC11, and these digests

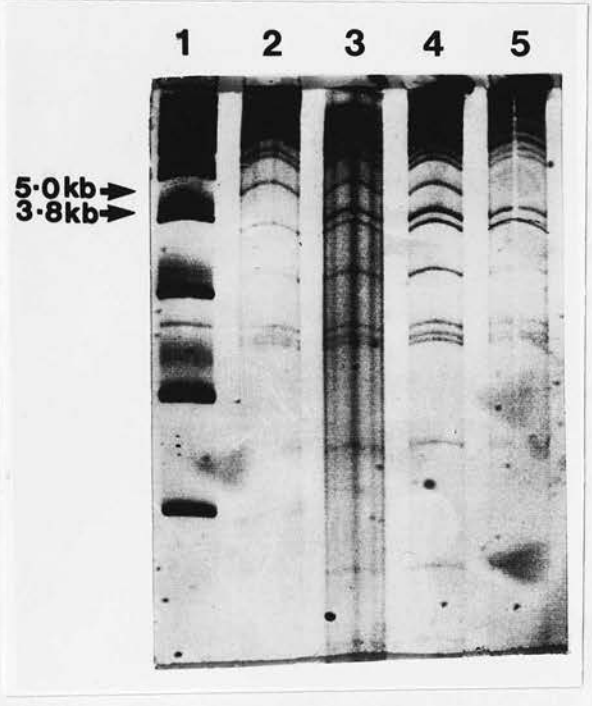
Figure 3.2

AHV-1 DERIVATIVES

Isolate	Derivative	Description
C500	PP	biologically cloned cell-associated virulent
C500	CA	cell-associated virulent
C500	CF	cell-free virulent
C500	CFA	cell-free attenuated
WC11	-	cell-free attenuated

Figure 3.3

Comparison of Sma I digests of PP, CA, CF and CFA C500 virus



Track 1 - 1kb ladder
Track 2 - PP digested with Sma I
Track 3 - CA digested with Sma I
Track 4 - CF digested with Sma I
Track 5 - CFA digested with Sma I
These samples were analysed using 7.5% acrylamide

were resolved alongside the original CFA preparation (CFA 1) and CA virus on 0.8% agarose. The profiles of these four samples were difficult to interpret on uv illumination of intercalated ethidium bromide therefore the gel was blotted and probed with DIG-labelled DNA from CA virus. The filter illustrated in figure 3.4 clearly demonstrates that the 5kbp fragment is absent from CFA 1, CFA 2 and WC11.

3.4 HOMOLOGY BETWEEN THE VARIABLE REGIONS

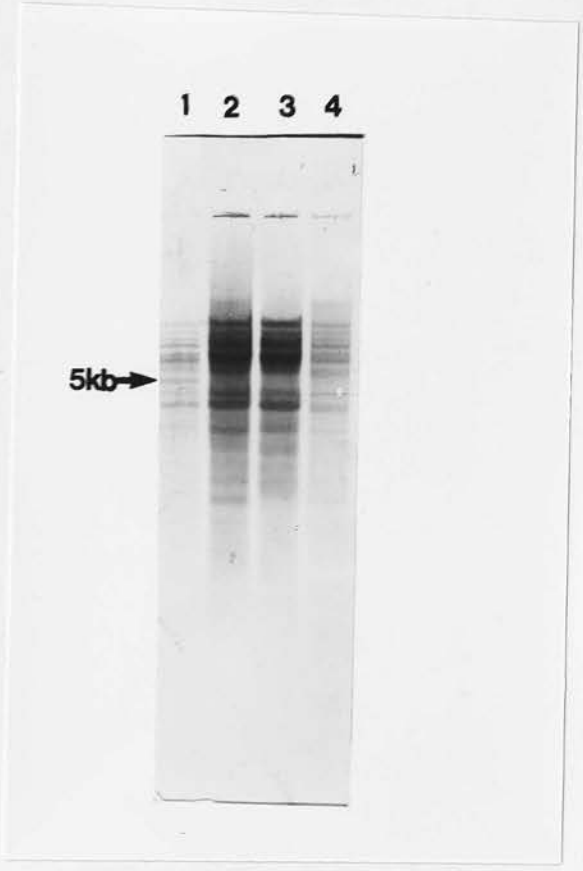
Digestion of the four C500 derivatives with Sma I revealed variation in the restriction profiles regarding fragments of 5 and 3.8kbp. The relationship between these fragments was initially examined by assessing whether the fragments were homologous. The 3.8kbp Sma I CFA fragment was recovered from agarose then labelled with DIG. The labelled DNA was used to probe a Southern blot of Sma I digests of PP, CA, CF and CFA genomic DNA (see figure 3.5), and homology was observed between the 5kbp PP Sma I fragment and the 3.8kbp CFA Sma I fragment. The 3.8kbp CFA probe hybridised to CA and CF fragments of 5 and 3.8kbp, a CFA fragment of 3.8kbp, and PP fragments of 5 and 8.6kbp. (The 8.6kbp PP fragment will be discussed further in chapter 5).

3.5 CLONING AND MAPPING OF THE 3.8KBP SMA I FRAGMENT

The 3.8kbp CFA Sma I fragment was excised from a 0.8% agarose gel and recovered using powdered glass. The vector selected to clone this fragment was the general purpose bacterial plasmid Bluescribe (pBS+), which was prepared by Sma I digestion and treatment with CIP. The 3.8kbp fragment was ligated into the prepared pBS+ and the resultant plasmid was transformed into JM101 bacterial cells then plated out onto L-agar containing X-gal, IPTG and L-amp. A single colony from this transformation was grown up and the plasmid extracted by the method described in 2.8.1. The resultant plasmid was digested with Sma I to produce an insert of 3.8kbp. This insert was excised from the plasmid and subsequently labelled with DIG. When used to probe a Sma I digest of CFA this probe hybridised to a single fragment of 3.8kbp, confirming that the correct fragment had been cloned. This clone was designated ATT-1, as it was the first clone to be derived from the attenuated C500 derivative.

Figure 3.4

Comparison of Sma I digests of attenuated AHV-1 variants

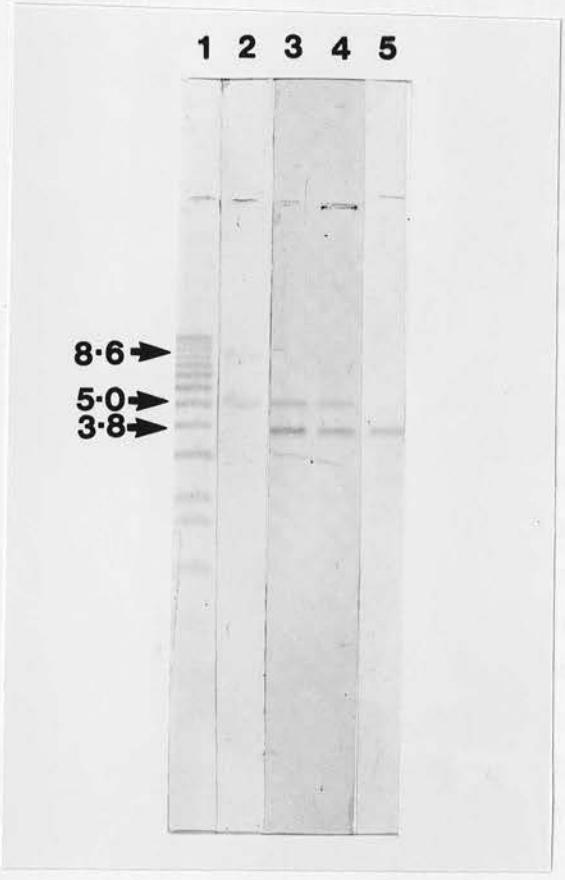


Track 1 - CA digested with Sma I
Track 2 - CFA1 digested with Sma I
Track 3 - CFA2 digested with Sma I
Track 4 - WC11 digested with Sma I

Tracks 1-4 were probed with DIG-labelled CA DNA

Figure 3.5

Sma I digests of C500 DNA probed with ATT-1



Track 1 - 1kb ladder
Track 2 - PP digested with Sma I
Track 3 - CA digested with Sma I
Track 4 - CF digested with Sma I
Track 5 - CFA digested with Sma I
Tracks 2-5 were probed with ATT-1 DNA labelled with DIG

The mapping of the ATT-1 clone was carried out by digestion with a series of REs which recognise unique or infrequent sites in pBS+. The details of how this mapping was achieved are described in appendix 2.1, and the resultant map is shown in figure 3.6.

3.6 CLONING AND MAPPING OF AN ATT-1 HOMOLOGUE FROM PP C500 VIRUS

The ATT-1 clone hybridised to a 5kbp Sma I fragment in the PP derivative. This 5kbp fragment was common to PP, CA and CF, therefore attempts were made to clone this fragment into pBS+. Repeated attempts to clone this fragment were unsuccessful therefore an alternative strategy was employed. PP DNA was digested with several other REs and the digests were resolved using 0.6% agarose, transferred to nylon and probed with the 5kbp fragment, which had been labelled with DIG. The 5kbp probe hybridised to a ladder of fragments when PP was digested with Kpn I or Pst I, but hybridised to a doublet of 3.6kbp in a Hind III digest. This doublet was recovered from agarose, ligated into pBS+, and used to transform JM101 bacterial cells. Three transformants were observed and one of these was grown overnight in L-agar with L-amp. The plasmid was extracted from this culture following the method described in 2.8.1 and the insert was isolated by digestion with Hind III, then labelled with DIG. This probe hybridised to PP Sma I fragments of 5 and 8.6kbp (the 8.6kbp fragment will be discussed further in chapter 5). This clone was designated VIR-1 as it was the first clone to be derived from virulent C500 virus.

The mapping of the VIR-1 clone was carried out by digestion with a series of RE with unique or infrequent recognition sites in pBS+. The experiments carried out to map this clone are described in appendix 2.2, and the resultant map of the VIR-1 clone is shown on figure 3.7.

3.7 THE DEGREE OF HOMOLOGY BETWEEN ATT-1 AND VIR-1

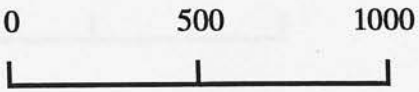
3.7.1 Introduction

The original 5kbp Sma I PP fragment had been shown to be homologous to the 3.8kbp Sma I CFA fragment. The two clones, VIR-1 and ATT-1, contained DNA from these homologous regions of PP and CFA respectively, and therefore were expected to exhibit a degree of homology. The termini of both inserts were

Figure 3.6

The completed restriction endonuclease map of ATT-1

Scale (in bp)



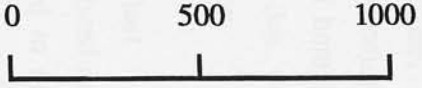
Key

H - Hind III
I - Hind II
K - Kpn I
L - Sph I
P - Pst I
S - Sma I
V - Pvu II

Figure 3.7

The completed restriction endonuclease map of VIR-1

Scale (in bp)



Key

C - Sac I
H - Hind III
I - Hind II
L - Sph I
P - Pst I
V - Pvu II

arbitrarily labelled 5' and 3'. Comparison of the RE profiles of the two clones suggested that homology might occur towards their 3' ends on the basis of fragment sizes. The 5' ends of the two clones differed with regard to both the recognition sites present and their locations.

3.7.2 Comparison of similar sized fragments by hybridisation

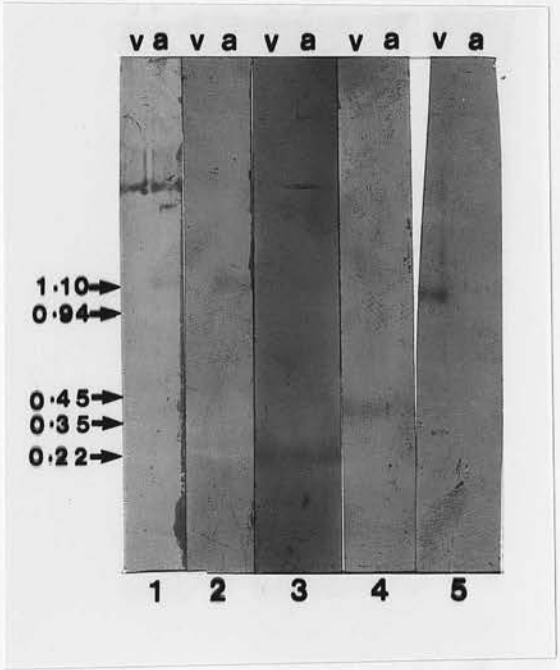
The nature of the derivation of the ATT-1 and VIR-1 clones suggested that a degree of homology existed between the two. This was confirmed when the ATT-1 insert, labelled with DIG, was used to probe the VIR-1 insert, and hybridisation was observed (data not shown). The degree of homology was initially assessed by comparing the restriction profiles of the two clones (see appendix 2 for fragment sizes). The two clones did not exhibit total homology as ATT-1 contained recognition sites for Kpn I, while VIR-1 did not. Conversely VIR-1 contained sites for Sac I, while ATT-1 did not. Several restriction fragments of identical size, however, were observed on digestion with Pst I/Hind III (fragment sizes of 435, 342 and 216bp) and with Pvu II (fragment size 844bp). These identically sized fragments were shown to be homologous by using the VIR-1 435, 342 and 216bp fragments (which are represented by fragments A, B and C respectively in figure 3.13) to probe Pst I/Eco RI digests of ATT-1 and VIR-1 (as a control). The fragments were labelled with DIG and homology was observed as illustrated in figure 3.8, filters labelled 2, 3 and 4. The remaining two filters in this figure were used to assess whether the DNA at the 3'-end of ATT-1 recognised sequences in VIR-1 (filter 1) and, conversely, to assess whether the DNA at the 5'-end of VIR-1 recognised sequences in the ATT-1 clone. In both instances no hybridisation was observed to the target DNA, demonstrating that the 5' and 3' ends of the clones differed. The Pst I/Hind III and Pvu II homologous fragments were aligned, as demonstrated in figure 3.9, suggesting that the minimum region of homology between the clones was approximately 1700bp.

3.7.3 Identification of the region where homology is lost

From the linear comparison of the two clones as illustrated in figure 3.9 the 5' ends of the two clones vary significantly with regard to the location of several restriction sites. This variation was thought to be too extensive to be due to point mutations and, as described in 3.7.2, the 5' end of the VIR-1 clone (fragment G in figure 3.13) was not homologous to any part of the ATT-1 clone. The lack of homology extended 3' to this Pst I fragment as fragment F in figure 3.13 also failed to exhibit homology to the ATT-1 clone, whilst conversely ATT-1 fragment

Figure 3.8

Probing digests of ATT-1 and VIR-1 with ATT-1 and VIR-1 subfragments

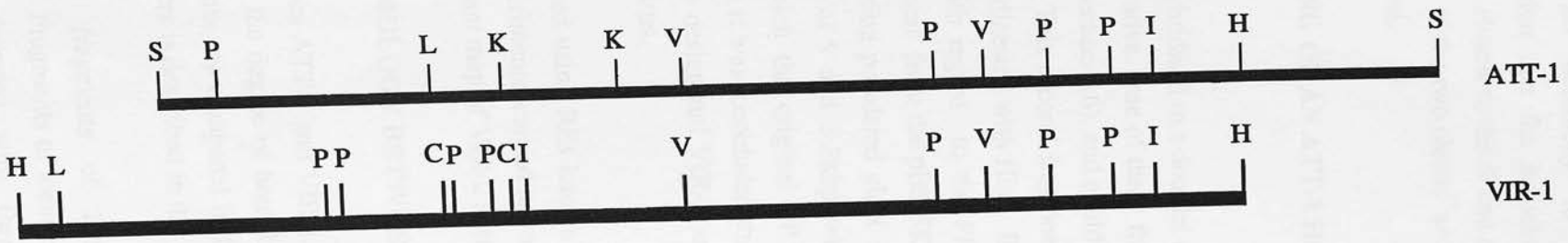


Tracks v - VIR-1 digests
Tracks a - ATT-1 digests

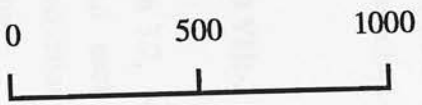
Probe 1 - figure 3.12, fragment A
Probe 2 - figure 3.13, fragment A
Probe 3 - figure 3.13, fragment B
Probe 4 - figure 3.13, fragment C
Probe 5 - figure 3.13, fragment G

Figure 3.9

Alignment of ATT-1 and VIR-1



Scale (in bp)



Key

C - Sac I	L - Sph I
H - Hind III	P - Pst I
I - Hind II	S - Sma I
K - Kpn I	V - Pvu II

H (figure 3.12) exhibited no homology to VIR-1. These hybridisation experiments led to the conclusion that the homology between the two clones is lost upstream of the Pvu II site closest to the 5' end of both clones. This homology loss is not due to misalignment of the two clones and suggests rearrangement of the genomic C500 DNA has occurred.

3.8 CLONING AND MAPPING OF AN ATT-1 HOMOLOGUE FROM CA C500 VIRUS

The 5kbp Sma I PP fragment hybridised to a doublet of 3.6kbp when used to probe a Hind III digest of the PP derivative. One of these fragments had been inserted into pBS+ and characterised (see section 3.6), and could therefore be distinguished from the second Hind III fragment. This second fragment was cloned into pBS SK+ after recovery from CA DNA digested with Hind III. (The relationship of the two 3.6kbp Hind III fragments with regard to the PP and CA derivatives will be discussed in chapter 5). The insert from the pBS SK+ clone was excised using Hind III, recovered from agarose using powdered glass then labelled with DIG. This probe hybridised to fragments of 5 and 3.8kbp when used to probe a Sma I digest of CA, the same fragments which the original PP Sma I 5kbp fragment recognised in this derivative, from which it was concluded that a homologous fragment had been cloned. This clone was designated VIR-2 as it was the second clone to be isolated from virulent C500 virus.

The VIR-2 clone was mapped using REs known to have sites in one or both of ATT-1 and VIR-1. The experiments carried out to map this clone are described in appendix 2.3, and the resultant map of VIR-2 is shown in figure 3.10.

3.9 THE DEGREE OF HOMOLOGY BETWEEN ATT-1, VIR-1 AND VIR-2

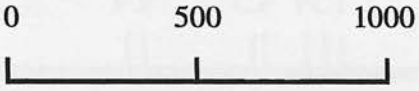
The known homology between ATT-1 and VIR-1, as described in section 3.7, was used as a basis to determine the degree of homology between VIR-2 and each of these clones. The three clones are compared in figure 3.11, and the derivation of the similarities presented here is described in the remainder of this section.

The VIR-2 clone produced fragments of 216bp, 342bp and 435bp when digested with Pst I/Hind III. Fragments of identical size were observed when both ATT-1 and VIR-1 were digested with these two enzymes. Hybridisation

Figure 3.10

The completed restriction endonuclease map of VIR-2

Scale (in bp)

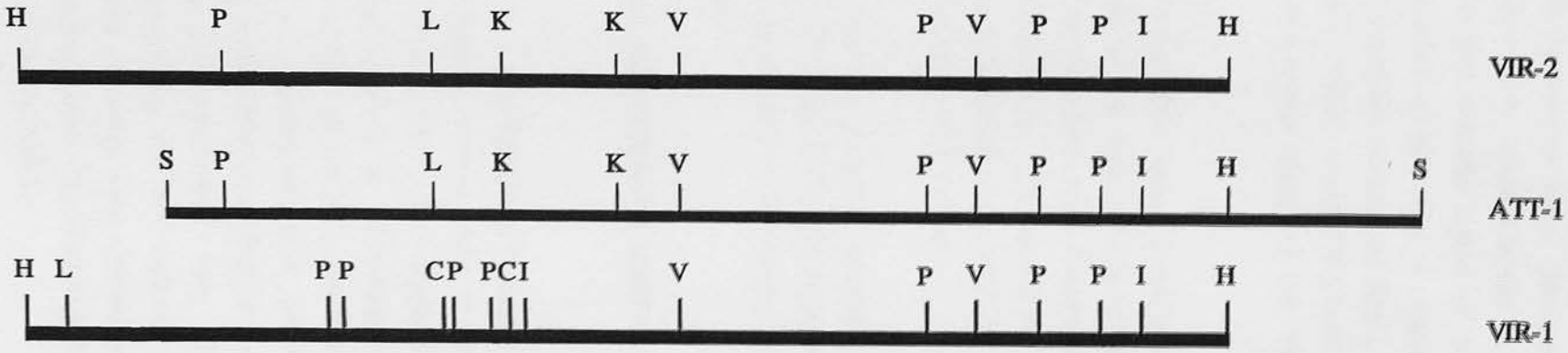


Key

H - Hind III
I - Hind II
K - Kpn I
L - Sph I
P - Pst I
V - Pvu II

Figure 3.11

Alignment of VIR-2, ATT-1 and VIR-1



Scale (in bp)

0 500 1000

Key

C - Sac I
H - Hind III
I - Hind II
K - Kpn I
L - Sph I
P - Pst I
S - Sma I
V - Pvu II

experiments showed these three fragments contained homologous DNA in all three clones. Pst I/Hind III digestion of VIR-2 also produced a larger Pst I fragment (fragment D in figure 3.16), identical in size to ATT-1 Pst I fragment E (figure 3.12) which was located over the region of homology loss between ATT-1 and VIR-1. Comparison of the Kpn I profiles of the three clones illustrated that while VIR-1 contained no sites for Kpn I, ATT-1 and VIR-2 both contained two sites for Kpn I, which resulted in a centrally located homologous fragment of 340bp. These comparisons suggested that VIR-2 was more similar to ATT-1 than to VIR-1.

The similarity between ATT-1 and VIR-2 appeared to span almost the entire ATT-1 clone 5' to the Hind III site as the Kpn I, Pst I and Sph I profiles exhibit (see figure 3.11). The 5'-ends of the two clones were found to differ in that ATT-1 was terminated by a Sma I site which had no equivalent in the VIR-2 clone, therefore the two clones were homologous upstream of the Hind III site in the ATT-1 clone until within 175bp of the 5' end of the ATT-1 clone.

The results described in the previous paragraph indicate that a far greater degree of similarity exists between ATT-1 and VIR-2 than exists with either of the above and VIR-1. The exact location of the sites of homology loss could only be achieved by sequencing the three clones.

3.10 VECTORS SELECTED FOR THE SUBCLONING OF ATT-1, VIR-1 AND VIR-2

The sequencing of ATT-1, VIR-1 and VIR-2 was facilitated by subcloning selected fragments into bacterial plasmid vectors. Prior to subcloning large scale preparations of ATT-1, VIR-1 and VIR-2 were prepared following the method described in 2.8.2. The vectors selected for subcloning were pBS+ and pBS SK+ as both contain sequences recognised by the M13 universal and reverse primers. These primers are orientated on opposing strands at either end of the multiple cloning sites of these vectors which therefore enables sequencing of both strands of the insert. The subclones were transformed into JM109 bacterial cells, amplified in L-broth containing 40μg ml^{-1} of L-amp and recovered following the method described in 2.8.1. The subclones were checked by labelling their inserts with DIG then using these DIG-labelled fragments to probe restriction digests of the relevant clone (i.e. ATT-1, VIR-1 or VIR-2).

3.11 PRIMERS USED TO COMPLETE THE SEQUENCING OF ATT-1, VIR-1 AND VIR-2

Subcloning of the three C500 clones did not enable sequencing of the entire inserts therefore oligonucleotides were used to complete gaps and orientate clones where appropriate. All oligonucleotides (Oswell DNA) were 20bp in length and their sequences are presented in appendix 3.

3.12 THE SUBCLONING AND SEQUENCING OF THE ATT-1 CLONE

3.12.1 The subcloning of ATT-1

The majority of the ATT-1 clone was subcloned using the restriction endonucleases (REs) Hind III, Pst I, Kpn I and Pvu II. The resultant subclones are demonstrated in figure 3.12a, denoted as fragments A-H. Initially the ATT-1 clone was digested with Pst I and fragments C, D and E were ligated into pBS+. Fragment B was isolated by digestion of ATT-1 with Pst I in conjunction with Hind III and inserted into pBS+ digested with the same two enzymes. Fragment A was subcloned by digesting ATT-1 with Hind III then self-ligating the larger fragment which corresponded to pBS+ ligated to fragment A at the 3' end via the Sma I site used in the original cloning of ATT-1. Fragments F and G were created by double digestion of the subclone containing fragment E with Pst I and Pvu II. The polylinker of pBS+ does not have a RE site for Pvu II, however this enzyme is a blunt-ended cutter, as is Sma I, which does have a site in the polylinker of pBS+. The two Pvu II/Pst I fragments were therefore cloned into pBS+ digested with Pst I and Sma I. The remaining two subclones, containing fragments H and I, were subcloned from a Kpn I digest of ATT-1. Fragment H was ligated into pBS+, and fragment I was cloned by self ligation of the ATT-1 Kpn I fragment which corresponded to pBS+ ligated to fragment I at the 5' end via the Sma I site used in the original cloning of ATT-1.

3.12.2 The sequencing of ATT-1

The subclones containing fragments A, B, C, D, F and H (as illustrated in figure 3.12a) were all completely sequenced using the M13 universal and reverse primers. The sequence of fragments A, B and F were achieved by resolving samples by electrophoresis of acrylamide sequencing gels for 8 hours using the "long mixes" supplied in the T7 sequencing kit. The orientation of fragments C and D could not be deduced from sequencing the relevant subclones because the fragments were

Figure 3.12

Figure 3.12a Map of the subclones of ATT-1 used for sequencing

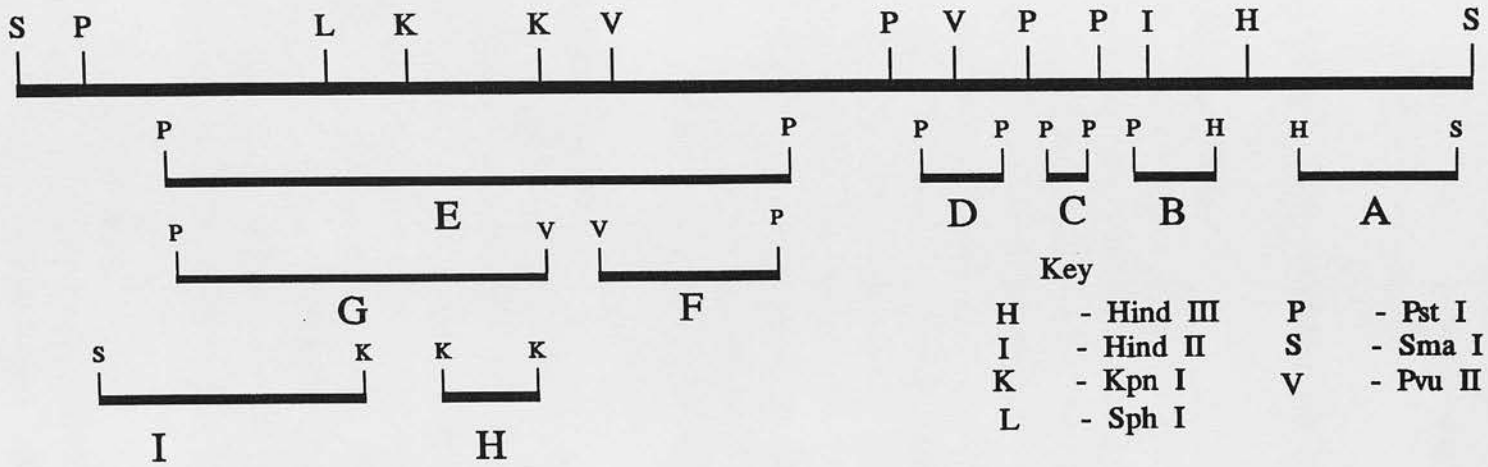
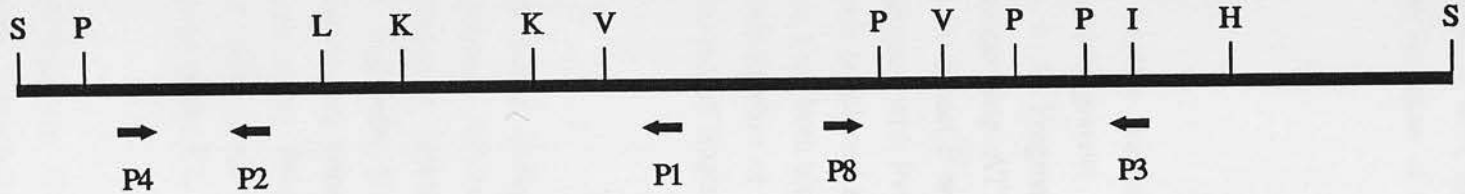


Figure 3.12b Location of oligonucleotides used in sequencing ATT-1



inserted into Pst I sites at both the 5' and 3' ends therefore primer 3 and primer 8 (figure 3.12b) were used to sequence from fragments B and E into this region. Primer 1 was used both to sequence the region between fragments F and H and to orientate fragment H. Fragment I was initially sequenced using the M13 universal and reverse primers, however, resolving samples in excess of 8 hours did not reveal the entire sequence of this clone therefore primer 2 and primer 4 were used as illustrated in figure 3.12b. The complete sequence and primer locations of the ATT-1 clone are presented in appendix 3.1.

3.13 THE SUBCLONING AND SEQUENCING OF VIR-1

3.13.1 The subcloning of VIR-1

The VIR-1 clone contained 7 sites for the RE Pst I therefore this enzyme was used in the first instance to subclone VIR-1 into appropriately sized fragments for sequencing. The fragments subcloned are illustrated in figure 3.13a. Fragments A, B and C were cloned in an identical manner to that of the corresponding ATT-1 clones, B, C and D, as described in 3.12.1. Pst I fragments D, E and F were subcloned into pBS+. The subclone containing fragment D was digested with Pvu II in conjunction with Pst I to yield fragments H and I, which were subcloned into pBS SK+ digested with Pst I and Sma I, since Pvu II and Sma I are both blunt-ended cutters. The remaining fragment, G, was subcloned by self-ligation of the largest fragment produced on Pst I digestion of VIR-1, as this consisted of fragment G ligated to pBS+ from the original cloning of VIR-1.

3.13.2 The sequencing of VIR-1

The subclones containing the inserts A, B, C, E, F, H and I (illustrated in figure 3.13a) were sequenced using the M13 universal and reverse primers. All clones were sequenced in both directions. Primers 3 and 8 were employed to orientate fragments B and C, while primers 5 and 7 were used to orientate fragments E and F. Fragment G was initially sequenced using the M13 universal and reverse primers, however, this region was too large to sequence using these primers alone. Primers 6 and 9 were employed to sequence the central region of this clone. The complete sequence and primer locations of VIR-1 are presented in appendix 3.2.

3.13.3 Sequence features of VIR-1

The location of primer 9 is 109bp from the Hind III terminus of fragment G and located immediately before a repeated region. This repeated region consisted of a

Figure 3.13

Figure 3.13a Map of the subclones of VIR-1 used for sequencing

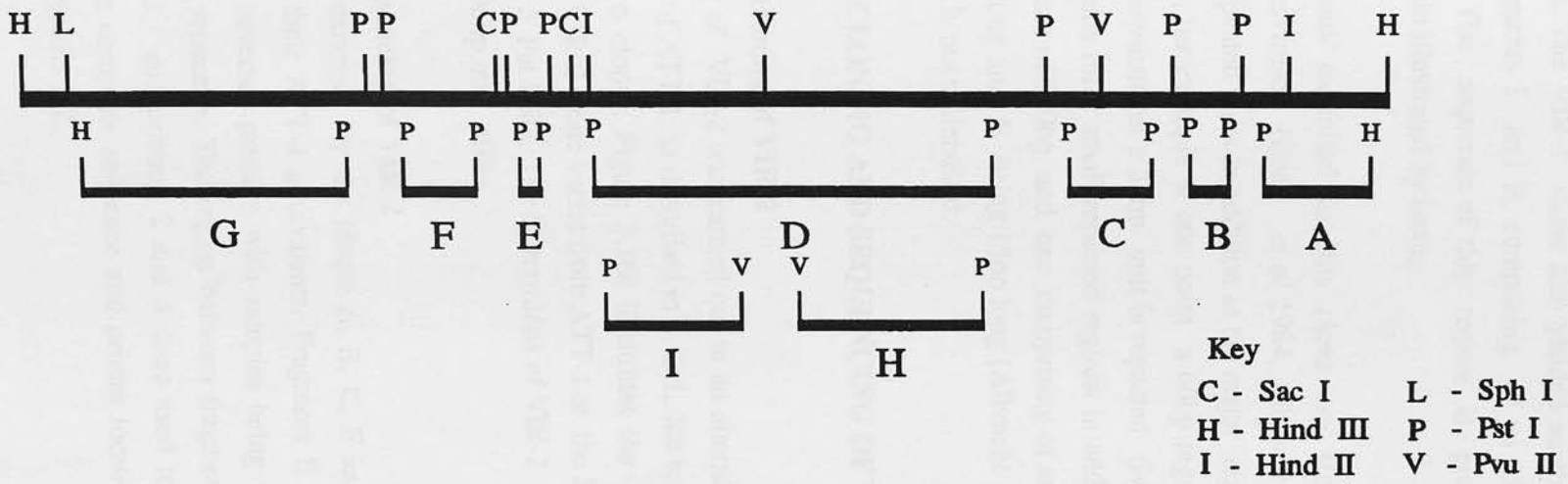
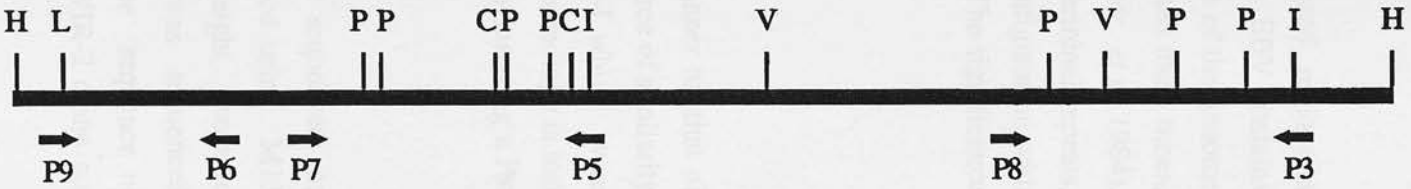


Figure 3.13b Location of oligonucleotides used in sequencing VIR-1



51bp sequence which was directly repeated five times, with an additional partial repeat of 26bp immediatedly upstream of the first copy. The sequence data of this region is presented in aligned form in figure 3.14, with minor differences indicated by boxed areas. The VIR-1 clone also yielded another region of repeated DNA, located in fragments I and E, comprising of an 85bp region which is directly repeated once. The sequence of this region is presented in figure 3.15, with differences again illustrated by boxes.

The repeat regions exhibited in this clone are similar to repeated regions in related γ-herpesviruses (Baer *et al* 1984, Albrecht *et al* 1992). EBV contains many small repeated units in addition to the major repeated regions of the genome (see figure 1.1), for example at one point a 60bp region is repeated three times, whilst further downstream a 33bp unit is repeated five times (Baer *et al* 1984). HVS also contains three small repeated regions in addition to the terminal repeats, one of 900bp, one of 550bp and one comprising of an ABABA configuration with A being 19bp long and B being 15bp long (Albrecht *et al* 1992). The significance of these repeats is not understood.

3.14 THE SUBCLONING AND SEQUENCING OF VIR-2

3.14.1 The subcloning of VIR-2

The subcloning of VIR-2 was carried out in an almost identical manner to that of the subcloning of ATT-1, as described in 3.12.1, due to the high degree of similarity between the two clones. Figure 3.16a illustrates the fragments A-I which were subcloned. The VIR-2 clone varies from ATT-1 at the 5' end in that homology is lost upstream of the 5' Pst I site. The 5' terminus of VIR-2 was cloned by inserting a Pst I fragment of 496bp into pBS+.

3.14.2 The sequencing of VIR-2

The subclones denoted by the letters A, B, C, F and H were all sequenced as described for their ATT-1 equivalents. Fragment E was sequenced using M13 universal and reverse primers with samples being resolved for eight hours to complete the sequence. The region between fragments F and H was sequenced using primer 1, and primers 2 and 4 were used to complete the sequence of fragment I. The complete sequence and primer locations of the VIR-2 clone are presented in appendix 3.2.

Figure 3.14

The repeated region of fragment G

Figure 3.16

Figure 3.16a Map of the subclones of VIR-2 used for sequencing

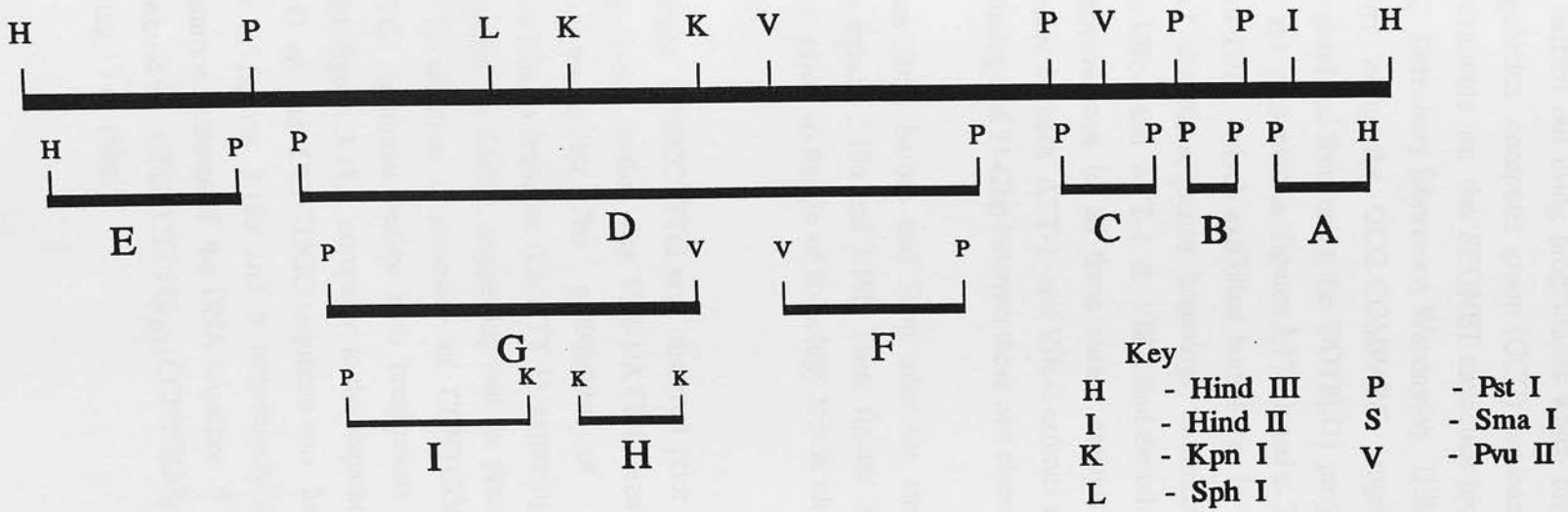
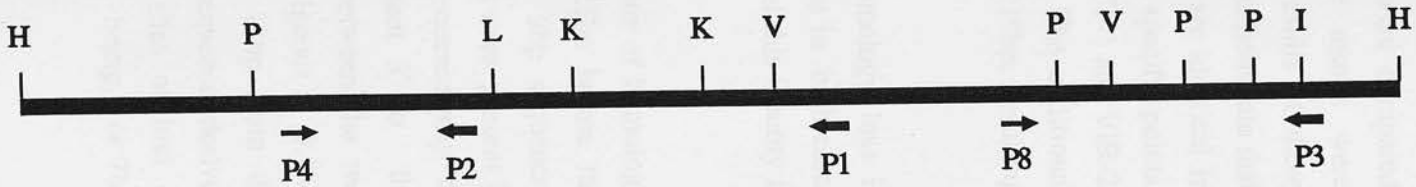


Figure 3.16b Location of oligonucleotides used in sequencing VIR-2



3.15 COMPARISON OF THE SEQUENCES OF ATT-1, VIR-1 AND VIR-2

Dot matrix analysis (Maizel and Lenk 1981) of the similarities between the three clones was carried out using programmes which form part of the University of Wisconsin genetics computer group (GCG) sequence analysis software package which was available on the SEQNET molecular biology computer facility based at the SERC Daresbury laboratory, Warrington, U.K. The clones were compared to each other using the GCG COMPARE programme, and the results were converted to graphical form using the DOTPLOT programme. The results of these programmes are presented as figures 3.17a, b and c. These figures demonstrate that the regions of DNA which exhibited homology by hybridisation are identical in sequence, and that the sequence homology is abruptly lost at two specific points, one between VIR-1 and ATT-1 & VIR-2 and the other between ATT-1 and VIR-2. The DNA homologous to all three clones extends for 1923bp. The additional sequence present in both ATT-1 and VIR-2 extends for a further 1219bp, resulting in a total homology of 3142bp between these two clones.

The sequence 50bp before and 50bp after the two sites of homology loss is illustrated in figures 3.18a and 3.18b. These figures demonstrate that in both cases the sequence prior to thesite of homology loss is identical and that this identity is abruptly lost.

A 9bp sequence, CTGCCTTTG was observed prior to the two sites of homology loss, ending 31bp before the VIR-1/ATT-1 breakpoint, and 25bp before the VIR-2/ATT-1 breakpoint. The probability of two identical 9bp sequences occurring in a 3.8kbp sequence (i.e. ATT-1), assuming all four bases were present in equal proportion, is 0.0002, suggesting that the presence of this sequence may be significant. In addition a sequence of CCNTGGG was present 3' to the CTGCCTTTG sequence before both breakpoints, with 7bp between the two sequences in figure 3.18a, compared to the sequences being contiguous in 3.18b. The final G of the CCNTGGG sequence was located 17 and 18bp from the breakpoints in figures 3.18a and b respectively. A consensus sequence derived from the common features of the DNA sequence 5' to the two sites of loss of homology would be $CTGCCTTTG_{(n)1}CCNTGGG_{(n)2}$, with (n)1 being 0 or 7bp, and (n)2 being 17 or 18bp.

Figure 3.17a Dotplot comparison of ATT-1 and VIR-1

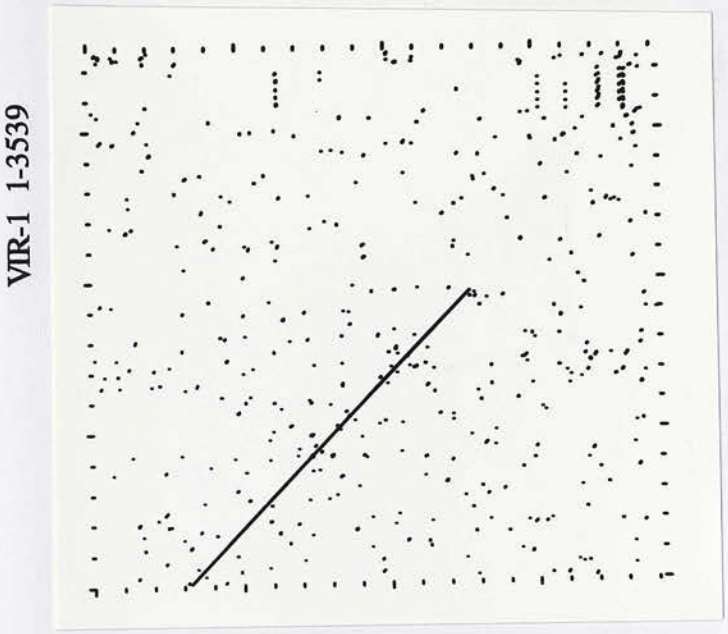


Figure 3.17b Dotplot comparison of ATT-1 and VIR-2

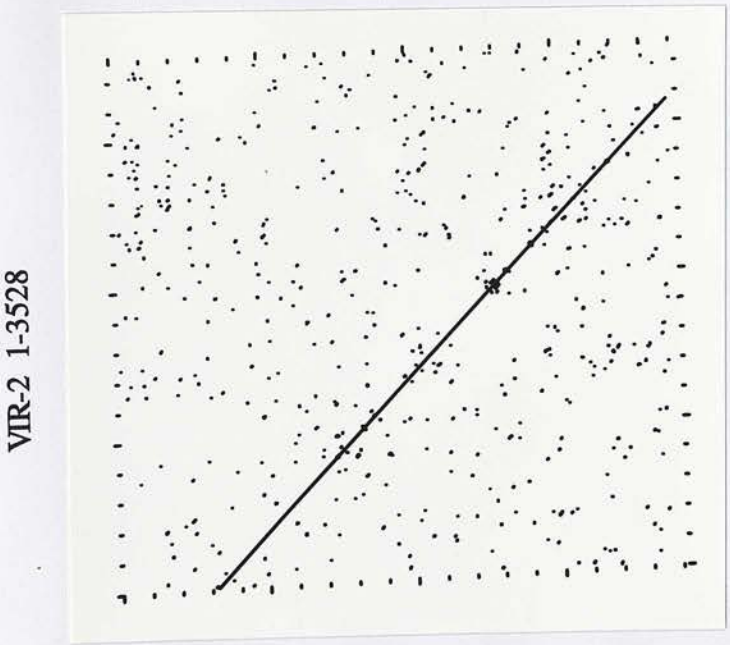


Figure 3.17c Dotplot comparison of VIR-1 and VIR-2

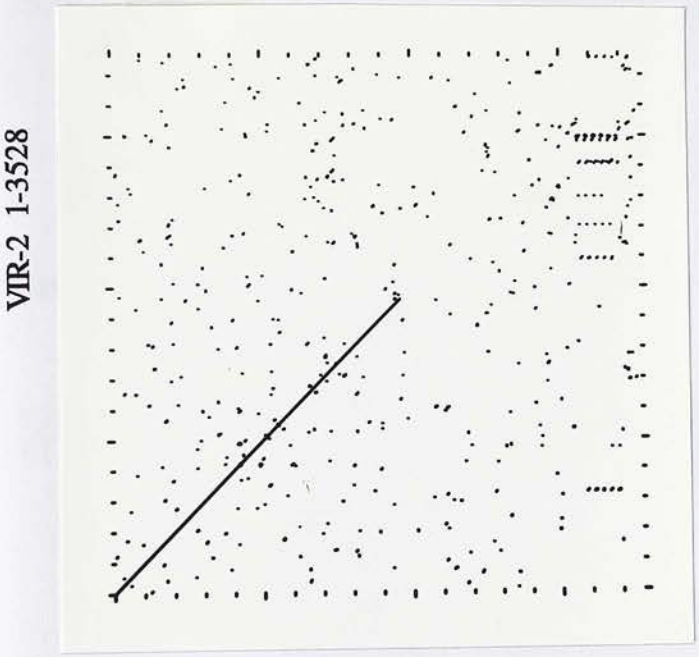


Figure 3.18

Figure 3.18a The site of homology loss between VIR-1 and ATT-1

Figure 3.18b The site of homology loss betwen VIR-2 and ATT-1

The sequences 3' to the sites of loss of homology bear little sequence similarity, although a pyrimidine rich region was observed in three of the four 3' sequences. The VIR-1 clone exhibited a sequence beginning 6bp 3' to the site of homology loss in which 18/21 residues were pyrimidines. ATT-1 did not contain an equivalent pyrimidine region 3' to this breakpoint. Pyrimidine rich regions were also observed 3' to the site of homology loss between VIR-2 and ATT-1. In VIR-2 a sequence beginning 28bp 3' from this site exhibited pyrimidines in 14/18 residues, while ATT-1 exhibited pyrimidines in 14/16 residues beginning 30bp 3' from this site.

The significance and function of these common sequence features located to either side of the sites of homology loss was not established. The presence of conserved sequences may indicate recognition sequences for functional proteins.

3.16 DISCUSSION

One aim of this project was to identify the genomic changes which are involved in the transition from virulence to attenuation. The experiments described in this chapter resulted in the identification of a region of the C500 genome which is variable between virulent and attenuated virus, although it is not necessarily the region of the genome responsible for virulence.

The 3.8kbp Sma I CFA fragment was cloned into pBS+. Attempts to clone the homologous PP 5kbp fragment were consistently unsuccessful. This failure may have been due to size constraints (as the proposed insert was 5kbp compared to the plasmid size of 3.2kbp), coupled with the fact that the cloning enzyme, Sma I, was a blunt-ended cutter. Linking fragments to generated restriction sites with overhangs to facilitate cloning was not attempted due to insufficient quantities of PP DNA (see chapter 5). Alternative explanations for the failure to clone this fragment included the possibility that the insert may encode and express a protein toxic to the bacterial cells which were used to amplify recombinant plasmids. A further possibility was that the fragment may contain repeated sequences which could result in an unusual conformation, hence the difficulties observed on cloning. The problems in cloning this 5kbp fragment were overcome by cloning two 3.6kbp Hind III fragments which exhibited homology to the 5kbp Sma I fragment.

The three clones, ATT-1, VIR-1 and VIR-2 were mapped with a range of REs and homology was assessed by comparison of the RE profiles and by hybridisation experiments. The three clones were shown to exhibit limited homologies to each other with the VIR-2 and ATT-1 clones being the most similar.

Sequencing the three clones revealed that the region of homology between VIR-1 and the other two clones is abruptly lost 1923bp from the 3'-terminus of this clone. This loss of homology is believed to be due to rearrangement occurring in the transition from virulence to attenuation. Similarly the homology between the ATT-1 and the VIR-2 clones is lost 3262bp from the 3'-terminus of the VIR-2 clone. The observation that comparison of the three clones exhibits two distinct sites of homology loss suggests that at least two genomic rearrangements occur during *in vitro* propagation of the C500 isolate of AHV-1.

The observation that the sequences 50bp before and after the two sites of homology loss exhibit common features suggests that a similar mechanism may be involved in creating these breakpoints.

Chapter 4
The location of the ATT-1 clone in the CFA genome

4.1 INTRODUCTION

The previous chapter of this thesis described the identification of a region of the C500 genome in which the arrangement of the DNA sequence differs between virulent and attenuated derivatives. Comparison of the two clones (VIR-1 and VIR-2) from virulent C500 virus and a clone (ATT-1) from attenuated C500 virus demonstrated two sites at which homology was lost, one between VIR-1 and VIR-2 and the other between VIR-2 and ATT-1. The events taking place at these regions of homology loss could be further understood by locating the three clones in the relevant C500 derivatives and by studying the regions adjacent to these clones. This chapter describes the experiments carried out to locate the ATT-1 clone in the CFA C500 genome and to examine the DNA adjacent to this clone. The WC11 laboratory isolate of AHV-1 had been studied in detail prior to the work undertaken in this thesis and a map of the majority of this genome was available (Bridgen 1991) therefore this was used as a starting point in the mapping studies described below.

4.2 HYBRIDISING ATT-1 TO WC11 RESTRICTION DIGESTS

The WC11 isolate of AHV-1 was originally isolated from a blue wildebeest. Continual serial passage of this AHV-1 isolate in tissue culture resulted in loss of virulence for both cattle and rabbits (Plowright *et al* 1965). Similarly, the CFA C500 derivative exhibited attenuation when used to infect rabbits (H.W. Reid, personal communication), therefore the CFA C500 genome was expected to be the most closely related C500 derivative to WC11. A map of WC11 for the restriction enzymes Hind III, Eco RI, Xho I, Sma I, Bam HI and Sal I is presented in figure 4.1 (Bridgen 1991).

WC11 DNA was digested with Bam HI, Sma I, Eco RI and Xho I respectively, with only Sma I cutting within the repeat region. These digests were resolved using 0.6% agarose, transferred to nylon filters then probed with DIG-labelled ATT-1 insert. The probe hybridised to a 10kbp Bam HI fragment, a 3.7kbp Sma I fragment, an Eco RI fragment of >15kbp and a Xho I fragment of >15kbp, as shown in figure 4.2. These results were difficult to interpret for three reasons. Firstly, the mapping of the WC11 genome was incomplete at the termini of the molecule therefore any of the above fragments may have been terminally located, particularly the 10kbp Bam HI fragment and the 3.7kbp Sma I fragment. In addition, Sma I digestion of WC11 resulted in three bands of approximately 3.7kbp therefore precise location of this Sma I fragment

Figure 4.1 Restriction map of WC11
(Incompletely mapped regions are indicated by hatched boxes)

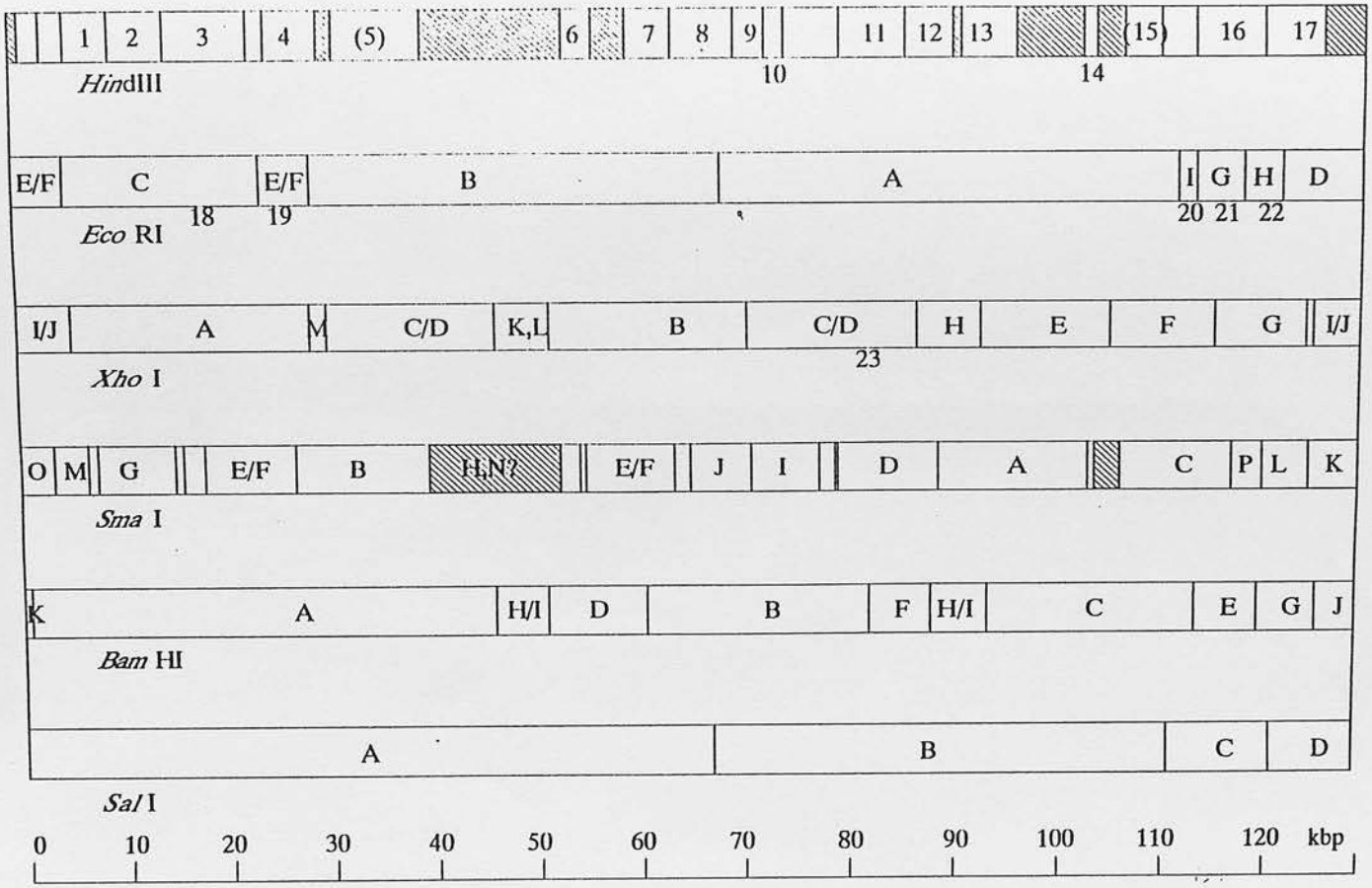
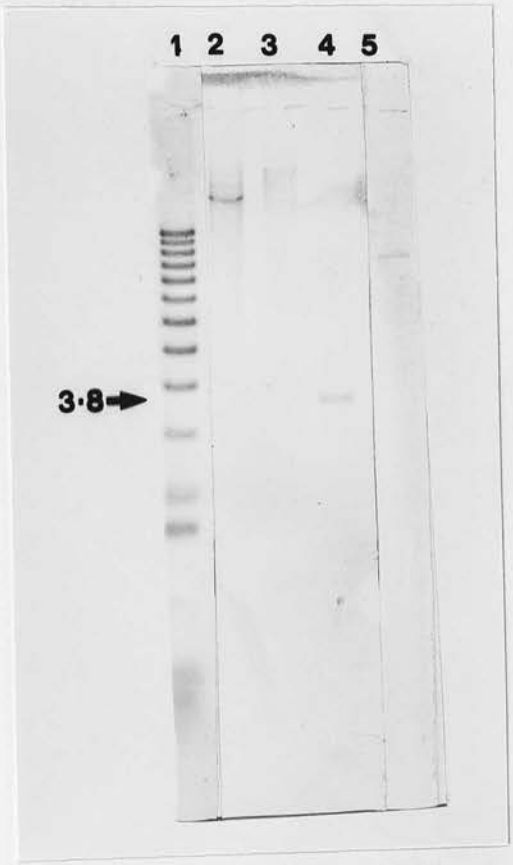


Figure 4.2

WC11 restriction digests probed with ATT-1



Track 1 - 1kb ladder
Track 2 - WC11 digested with Eco R1
Track 3 - WC11 digested with Xho I
Track 4 - WC11 digested with Sma I
Track 5 - WC11 digested with Bam HI

Tracks 2-5 were probed with ATT-1 labelled with DIG

recognised by ATT-1 was not achieved on the basis of this analysis. Finally, the results of probing Eco RI and Xho I WC11 digests with ATT-1 were inconclusive as the sizes of the fragments recognised by the probe were greater than the largest of the molecular markers used.

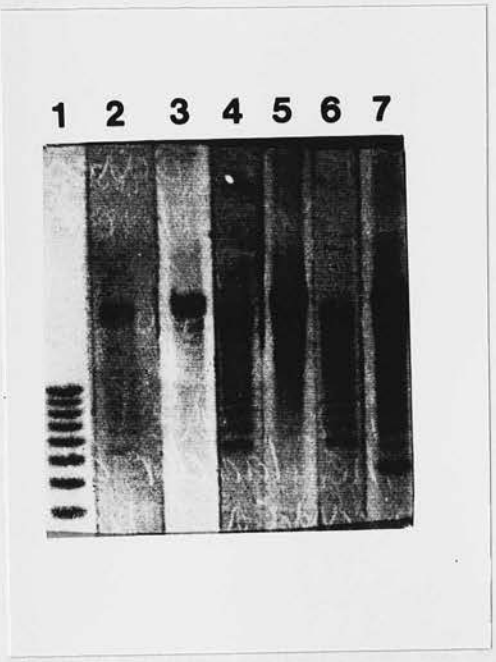
4.3 COMPARATIVE HYBRIDISATIONS OF ATT-1 TO C500 RESTRICTION DIGESTS

The results described above did not locate ATT-1 in the WC11 genome. The ATT-1 clone was derived from the C500 CFA derivative, which, like WC11, was a cell-free attenuated AHV-1 isolate. ATT-1 was used to probe Bam HI, Eco RI and Xho I CFA digests respectively to compare the ATT-1 hybridisation profiles of the WC11 and CFA genomes. In addition, the virulent PP C500 derivative (see figure 3.2) was also digested with these three enzymes and probed with DIG-labelled ATT-1 insert to determine whether ATT-1 would hybridise to similar sized fragments in virulent and attenuated C500 virus. These hybridisations illustrated that with both C500 derivatives digestion with each enzyme and subsequent probing with ATT-1 produced a ladder of fragments (see figure 4.3). The PP and CFA profiles were not identical with any of the three REs e.g. the smallest PP Bam HI fragment was approximately 7.5kbp, whilst the smallest CFA Bam HI fragment was approximately 6kbp. These differences in fragment sizes suggested that the PP and CFA genomes were arranged differently in the region which is recognised by ATT-1.

The WC11 and C500 profiles on digestion with Bam HI, Eco RI and Xho I respectively were distinct in that C500 exhibited a laddering effect while WC11 did not. This suggested that the location of ATT-1 differed in the genomes of these two AHV-1 isolates. The laddering effect suggests that the ATT-1 clone was terminally located in C500, since Bam HI, Eco RI and Xho I recognition sequences were absent from the repeat unit of this isolate. The WC11 repeat unit did not contain recognition sites for any of these three enzymes (Bridgen *et al* 1989), however, a similar laddering effect was not observed in the WC11 tracks, suggesting that ATT-1 was located further from the terminal repeats in WC11 than in C500. As a result of these observations, further mapping studies were restricted to the C500 isolate.

Figure 4.3

PP and CFA restriction digests probed with ATT-1



Track 1 - 1kb ladder
Track 2 - PP digested with Eco R1
Track 3 - CFA digested with Eco RI
Track 4 - PP digested with Xho I
Track 5 - CFA digested with Xho I
Track 6 - PP digested with Bam HI
Track 7 - CFA digested with Bam HI
Tracks 2-7 were probed with DIG-labelled ATT-1

4.4 HYBRIDISATION OF ATT-1 TO THE C500 CFA DERIVATIVE

The DIG-labelled ATT-1 insert was used to probe a Hind III digest of CFA DNA, as the recognition site for Hind III was also absent from the C500 repeat unit. A laddering effect was observed as illustrated in figure 4.4a. When this laddering effect was more closely examined each fragment appeared to have another fragment directly above it, producing the overall effect of a double ladder.

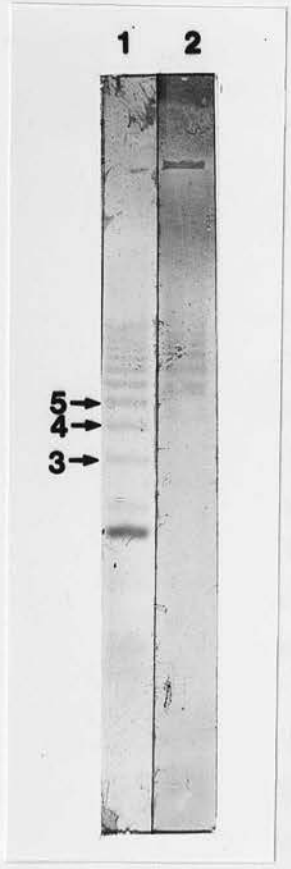
The ATT-1 clone contained an internal recognition site for Hind III, as illustrated in figure 4.4c. This Hind III site cut the ATT-1 insert into two fragments, one of approximately 3.2kbp (fragment A) and the other of approximately 0.6kbp (fragment B). The ATT-1 clone was digested with Hind III and fragments A and B were subcloned into pBS+ in order to prevent contamination of one fragment with another during probe preparation. The inserts of these subclones were labelled with DIG and used to probe CFA DNA digested with Hind III. These hybridisations demonstrated that a laddering effect was present when each of the probes was used individually to probe CFA Hind III digests, as shown in figure 4.4b. These ladders were distinct as the smallest fragment observed using fragment A as probe was 4.2kbp, whilst the smallest fragment observed using fragment B was 3.9kbp. This difference in the individual ladders accounted for the double laddering effect observed when the entire ATT-1 insert was used as probe.

The Hind III recognition sequence was not present in the repeat unit of C500, therefore sequence located before the first Hind III site of the unique region would be associated with fragments of varying sizes depending on the number of repeat units attached to the genomes of the individual virions. This virion variation resulted in the laddering effect observed, implying that ATT-1 was terminally located and suggesting that the internal Hind III site of the ATT-1 clone was the first Hind III site of the unique region of the CFA genome. On this basis either fragment A or fragment B was expected to produce a laddering effect dependent on the orientation of ATT-1 with respect to the terminal repeats. One possible explanation for the observation that both fragments A and B produced a laddering effect when used to probe a CFA Hind III digest was that one or both of these fragments recognised sequence located within the terminal repeat unit.

Figure 4.4

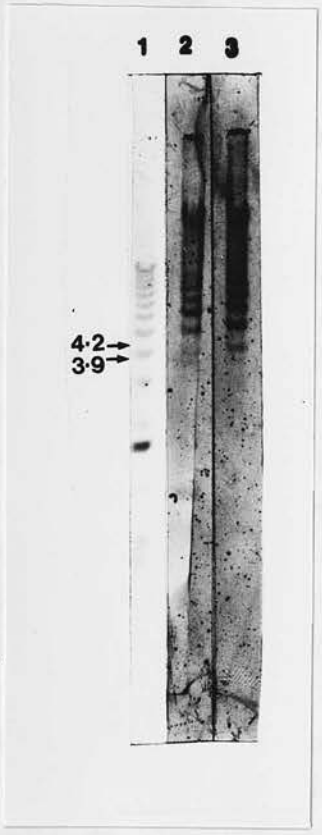
CFA Hind III digests probed with ATT-1 and the Hind III fragments of ATT-

Figure 4.4a



Track 1 - 1kb ladder
Track 2 - CFA digested with Hind III and probed with ATT-1 labelled with DIG

Figure 4.4b



Track 1 - 1kb ladder
Track 2 - CFA digested with Hind III and probed with fragment B
Track 3 - CFA digested with Hind III and probed with fragment A

Figure 4.4c Hind III map of ATT-1

4.5 CLONING OF THE REPEAT UNIT OF C500

The repeat unit of C500 was known to be represented by a supermolar fragment of 1050bp on digestion of the CF C500 derivative with either Hind II, Mbo I or Sac I (Bridgen *et al* 1989). This supermolar fragment was subsequently observed on Hind II and Sac I digestion of the three C500 derivatives, PP, CA and CFA, employed in this study. Genomic CFA C500 DNA was digested with Sac I and the 1050bp fragment was recovered from agarose then ligated into the Sac I site of pBS+. This clone was designated CR-1.

A probe of the CR-1 insert was prepared by Sac I digestion of the clone prior to labelling with DIG. Two experiments were undertaken to establish whether CR-1 contained the 1050bp C500 repeat unit. Firstly, the CR-1 clone was compared to the WC11 repeat clone, M30 (Bridgen *et al* 1989), which contained the WC11 Hind II 950bp repeat unit ligated into pBS+. M30 was digested with Pst I in combination with Eco RI to isolate the insert. This was then probed with DIG-labelled CR-1 insert and homology was observed. Secondly, the CR-1 DIG-labelled insert was used to probe PP, CA and CFA Bam HI digests and a laddering effect was observed in all tracks (data not shown). These two results indicated that the CR-1 insert was the C500 1050bp repeat unit and that the desired DNA fragment had been cloned.

The CR-1 clone was further characterised and compared to M30 by Sma I digestion of the two clones. The two inserts were each isolated by electrophoresis using 1.0% agarose, following double digestion with Pst I/Eco RI. The inserts were recovered using powdered glass, digested with Sma I then resolved using 7.5% polyacrylamide. The M30 insert produced bands of 170, 260, 280 and 320bp, while the CR-1 clone produced bands of 180, 290, 300 and 320bp (data not shown). The frequency of Sma I sites in both clones suggested that both clones were G+C rich.

4.6 COMPARISON OF ATT-1 TO CR-1

A DIG-labelled probe was prepared from the CR-1 insert and used to probe a Pst I digest of ATT-1 which had been resolved using 1.0% agarose and transferred to nylon. No hybridisation was observed between the CR-1 Sac I insert and the ATT-1 Pst I fragments. This lack of homology indicated that the laddering effect observed in section 4.4 was not due to ATT-1 sharing homology with the C500 repeat region, and therefore must be due to ATT-1 being located adjacent to the repeats.

4.7 POSSIBLE LOCATIONS OF ATT-1 ON THE CFA GENOME

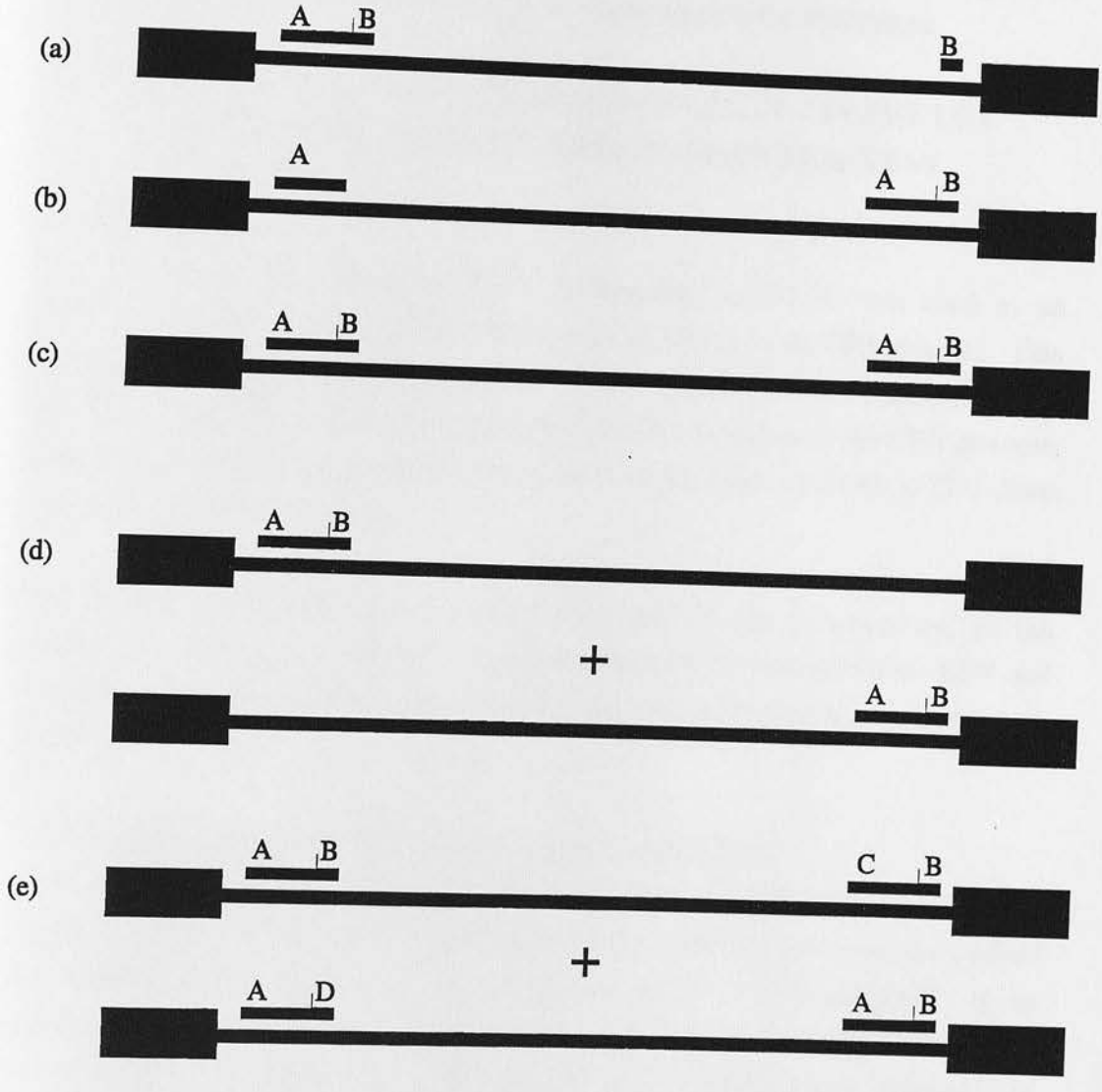
ATT-1 subclones A and B produced distinct laddering effects when used to probe Hind III CFA digests. This was not due to homology between ATT-1 and the repeat unit of C500. Initially, the laddering effect produced on probing Hind III CFA digests with ATT-1 was interpreted by assuming that either fragment A or fragment B represented unique DNA located between the terminal repeats and the first Hind III site of the unique DNA. This interpretation was proved incorrect by the observations that fragments A and B independently produced a laddering effect when used to probe Hind III digests of CFA DNA. An alternative explanation for the presence of these ladders was that all or part of the ATT-1 clone was repeated such that it was present in the same orientation at both ends of the CFA genome.

The five possible genomic arrangements to explain these ladders are illustrated in figure 4.5. The possible arrangements are:- (a) the complete ATT-1 clone is present at the left-hand side of the genome and in addition B, or part of B is also present at the right-hand side of the genome; (b) the complete ATT-1 clone is present at the right-hand side of the genome and in addition A, or part of A is present at the left-hand side of the genome; (c) the complete ATT-1 clone is present at both ends of the genome in the same orientation; (d) CFA consists of a mixed population whereby two genomic arrangements are present, one with the ATT-1 clone at the 5' end and one with the ATT-1 clone at the 3' end; (e) either fragment A or fragment B is repeated in the genome, once as ATT-1 and once with another fragment, of identical size to either A or B respectively.

The arrangements demonstrated in figures 4.5a and b were eliminated by the fact that when ATT-1 was used to probe a CFA Sma I digest (see figure 3.5) only one fragment of 3.8kbp was observed, with no hybridisation to fragments of different sizes to represent either A or B alone. The arrangement in 4.5d appeared less likely than that presented in 4.5c due to the observation that when a CFA Sma I digest was probed with DIG-labelled CA DNA a supermolar 3.8kbp fragment was observed (see figure 3.4), suggesting that this fragment was present twice. Alternatively this supermolar fragment may have been due to the presence of more than one 3.8kbp Sma I fragment, particularly since WC11 contains three Sma I fragments of approximately this size. The presence of more than one 3.8kbp Sma I fragment introduces the possibilities outlined in 4.5e, however the chances of fragment C

Figure 4.5

Possible arrangements of ATT-1 in the CFA genome



Fragment A = 3.2kbp
Fragment B = 0.6kbp
Fragment C = 3.2kbp ≠ Fragment A
Fragment D = 0.6kbp ≠ Fragment B

being identical in size to fragment A, and likewise fragment D being identical in size to fragment B seemed unlikely.

The simplest working model from the possible genomic arrangements illustrated in figure 4.5 was 4.5c, i.e. that the ATT-1 clone is present at either end of the genome in the same orientation. In order to determine which of the possible arrangements outlined in figure 4.5 was correct additional experiments were undertaken.

4.8 INVESTIGATION OF THE LOCATION OF ATT-1 IN THE CFA GENOME USING THE POLYMERASE CHAIN REACTION

4.8.1 Introduction

The polymerase chain reaction (PCR), as described in 2.8.20, was used as an alternative to RE data to investigate the location of ATT-1 in the CFA genome. This technique was used to attempt to amplify the presumptive junction regions between the ends of the ATT-1 clone and the repeats at either terminus of the CFA genome. The primers for these reactions were designed from the sequence of the ATT-1 clone and the CR-1 clone (see 4.8.2).

The termini of the CFA genome were designated 5' and 3' arbitrarily, as the orientation of the genome relative to the related sequenced γ-herpesviruses EBV and HVS has not been established with regard to the C500 sequence represented by the ATT-1 clone.

4.8.2 Sequencing and selection of PCR primers from CR-1

Ends of the CR-1 clone were sequenced in order to generate sufficient data to design oligonucleotides for use in PCR. The sequencing was achieved following the method described in 2.8.21, involving Taq polymerase and deoxy-7-deaza-dGTP, in an attempt to prevent secondary structures caused by the high G+C content of the repeat unit (71% G+C as assessed by caesium chloride density centrifugation [Herring *et al* 1983]). The limited sequence presented in figure 4.6 was generated from only one strand of DNA, due to the presence of secondary structures. Figure 4.6 shows the Sac I site in a central location due to orientation and amalgamation of the sequence generated from the 5' and the 3' termini of this clone. (This arrangement represents the two sequences as contiguous as this is how they would be arranged *in vivo*). The primers selected for PCR are outlined in figure 4.6 and denoted RP1 and RP2.

Figure 4.6

Sequence data from the CR-1 clone

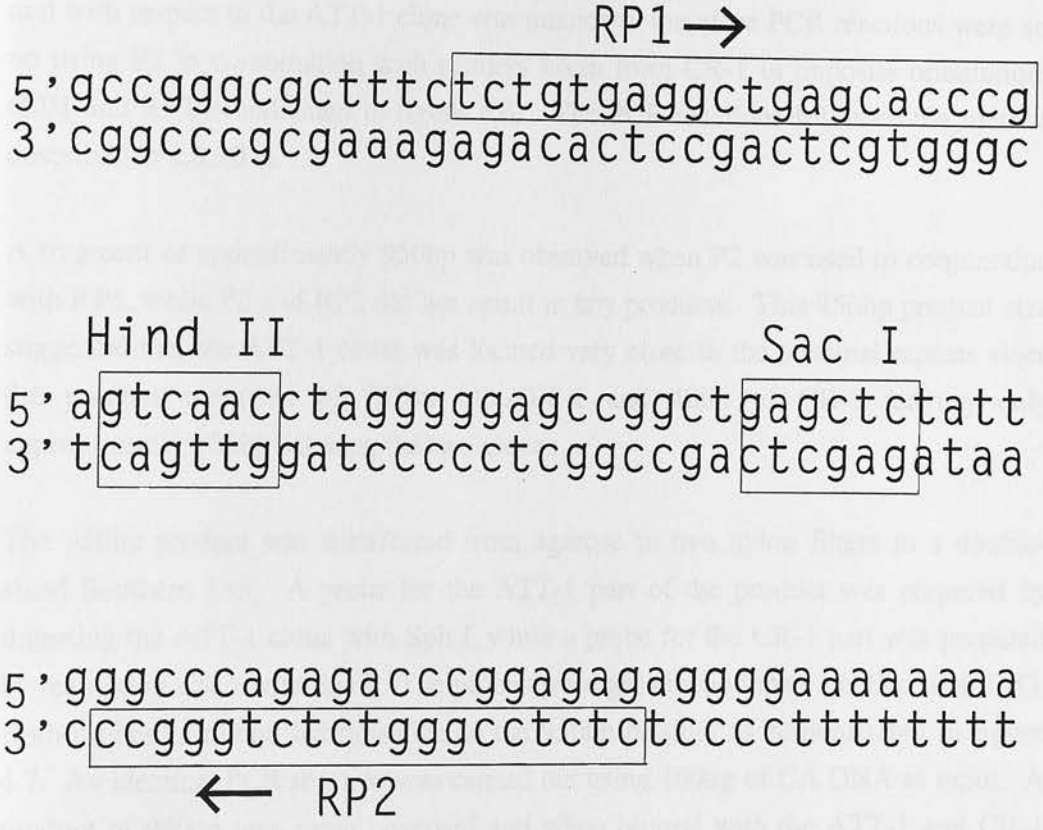


Figure 4.6 also demonstrates the location of the Hind II site which was known to be present in the C500 repeats. This site is located 24bp from the Sac I site.

4.8.3 The use of PCR to determine whether the 5' terminus of the ATT-1 clone is adjacent to the C500 terminal repeats

The oligonucleotides designed for sequencing the ATT-1 clone were used for PCR. In the first instance P2, located 725bp from the 5' end of the ATT-1 clone (see appendix 3 and figure 4.9) was used in conjunction with primers from the repeat region. The orientation of the CR-1 clone relative to its position in the CFA genome and with respect to the ATT-1 clone was unknown, therefore PCR reactions were set up using P2 in combination with primers taken from CR-1 in opposite orientations (RP1 and RP2 as indicated in figure 4.6). The PCR reactions followed the method described in 2.8.20 A.

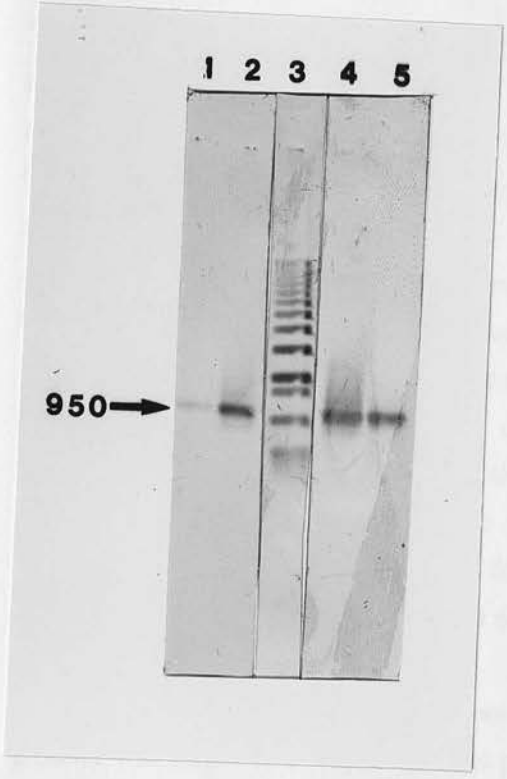
A fragment of approximately 950bp was observed when P2 was used in conjunction with RP1, while P2 and RP2 did not result in any products. This 950bp product size suggested that the ATT-1 clone was located very close to the terminal repeats since the product consisted of 752bp of ATT-1 and 48bp of CR-1 leaving only approximately 150bp between the two clones.

The 950bp product was transferred from agarose to two nylon filters in a double-sided Southern blot. A probe for the ATT-1 part of the product was prepared by digesting the ATT-1 clone with Sph I, while a probe for the CR-1 part was prepared by restricting CR-1 with Sac I. The probes were gel purified then labelled with DIG. Both probes exhibited hybridisation to the 950bp product as demonstrated in figure 4.7. An identical PCR reaction was carried out using 100ng of CA DNA as input. A product of 950bp was again observed and when blotted with the ATT-1 and CR-1 probes (as for the CFA product) hybridisation was observed. This result is illustrated in tracks 2 and 4 in figure 4.7, and is discussed in section 5.9.

An additional PCR was carried out using a primer further from the 5' terminus of the ATT-1 clone to verify the location of this clone as being adjacent to the terminal repeats. The primer selected was located 1564bp from the 5' terminus of the clone (P1 as illustrated in appendix 3.1). This primer was used in conjunction with RP1 following the conditions described in 2.8.20 B, with the target being CFA DNA. The predicted size of product was approx 1750bp (1564bp from ATT-1, 150bp from the sequence between the clones and 48bp from CR-1). The PCR products were

Figure 4.7

Probing PCR products from RP1 and P2



Track 1 - CFA RP1/P2 product
Track 2 - CA RP1/P2 product
Track 3 - 1kb ladder
Track 4 - CA RP1/P2 product
Track 5 - CFA RP1/P2 product

Tracks 1 and 2 were probed with an ATT-1 fragment labelled with DIG
Tracks 4 and 5 were probed with CR-1 labelled with DIG

resolved through 0.8% agarose and a single fragment of 1700bp was observed (data not shown). Similarly a PCR reaction was set up using P3 and RP1 following the reaction conditions described in 2.8.20 A. The predicted product size was approximately 3kbp. A fragment of 3kbp was observed (see figure 4.8), in addition to submolar fragments of varying sizes. These submolar fragments were believed to be artefacts of PCR.

The PCR reactions using primer combinations of RP1 with P1, P2 and P3 respectively suggested that fragment A in its entirety was located adjacent to the repeats at the 5' end of the C500 CFA genome. In order to determine whether fragment B was located adjacent to fragment A at the 5' end of the genome a PCR reaction was set up using primer 11 and RP1. The reaction conditions were as described in 2.8.20 C, with the predicted fragment size being approximately 3.4kbp but, when the completed reaction was analysed, no products were observed. The reaction was repeated twice, with a second aliquot of primers, dNTPs, reaction buffer and Taq polymerase added after 15 cycles to allow for reagent depletion. The reaction conditions were as described above, but the initial extension time was raised to 4 minutes in one variation, and to 10 minutes (with no extension per cycle) in another. No products were observed following either set of conditions. This result was inconclusive as no equivalent size control was available.

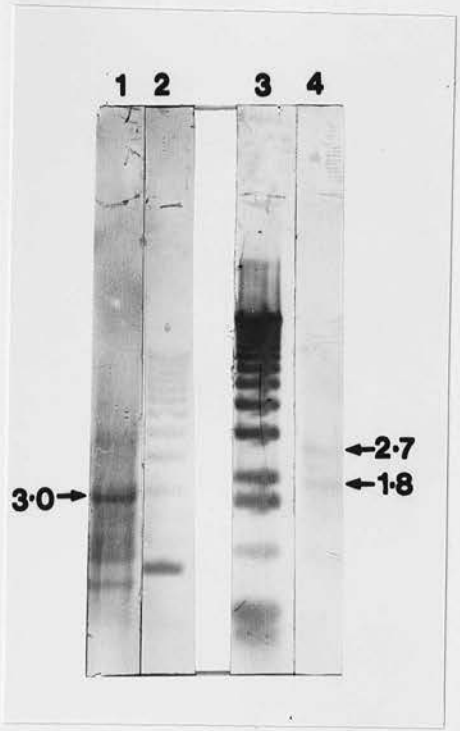
The lack of products observed may have been due to size constraints as the desired product was larger than any fragments previously amplified in this study. The M13 universal and reverse primers were used to attempt to amplify the 3.8kbp entire ATT-1 insert as a size control, but no 3.8kbp product was observed. Further location studies with regard to the orientation of fragment B at the 5' end of the CFA genome were restricted to probing Southern blots of genomic digests.

4.8.4 The use of PCR to determine whether the 3' terminus of the ATT-1 clone is adjacent to the C500 terminal repeats.

The results described in 4.8.3 determined that fragment A was adjacent to the 5' repeats. The genomic arrangement illustrated by figure 4.5c suggested that ATT-1 was located adjacent to the repeats at either end of the genome. This was further investigated by a PCR reaction carried out using a primer orientated towards the 3' terminus in combination with RP2 from the repeat unit. In the first instance the primer selected was P12, located 33bp from the 3' terminus of ATT-1 (see appendix 3.1 and figure 4.9). The reaction was carried out using 100ng of CFA DNA

Figure 4.8

PCR products from CFA DNA



Track 1 - CFA product from P3/RP1 probed with ATT-1
Track 2 - 1kb ladder
Track 3 - 1kb ladder
Track 4 - CFA product from P14/RP2 probed with ATT-1 fragment B

following the conditions described in 2.8.20 A. The product resulting from this reaction was approximately 500bp in length (see figure 5.14), with 33bp of this from ATT-1 and 30bp from CR-1. This result indicated that the ATT-1 clone was located very close to the terminal repeats at the 3' end of the genome. This location of ATT-1 was further suggested by a PCR reaction involving P14 and RP2. P14 was located within fragment A, 1390bp from the end of the ATT-1 clone. The reaction conditions were as described in 2.8.20 D. Two products were observed, one of 1.8kbp and another of 2.7kbp. Both of these fragments exhibited homology to ATT-1 (see figure 4.8). The 2.7kbp fragment was believed to represent the region between ATT-1 and the terminus of the unique region with one repeat unit also present, which would result in a predicted fragment size of 1.8 + 1.05bp.

The two PCR reactions described above suggested that ATT-1 fragment B was adjacent to the 3' terminal repeats, and that fragment A was adjacent to this location of fragment B.

A further PCR reaction was attempted using P4 and RP2 to determine whether all of fragment A was adjacent to fragment B at the 3' terminus of the CFA genome. The reaction conditions were as described for P11 and RP1, with a second aliquot of reagents added after 15 cycles. The predicted size of the PCR product from this reaction was approximately 4.1kbp. Although repeated attempts to generate product from this reaction were consistently unsuccessful this result was not conclusive due to size constraints in the PCR.

4.8.5 Conclusions

The PCR reactions carried out in this section are summarised in figure 4.9. Fragment A was established as being adjacent to the 5' terminal repeats, but fragment B was not confirmed as being adjacent to this 5' copy of fragment A. Fragment A was established as being adjacent to fragment B at the 3' terminus of the molecule, and in addition this copy of ATT-1 was demonstrated to be terminally located.

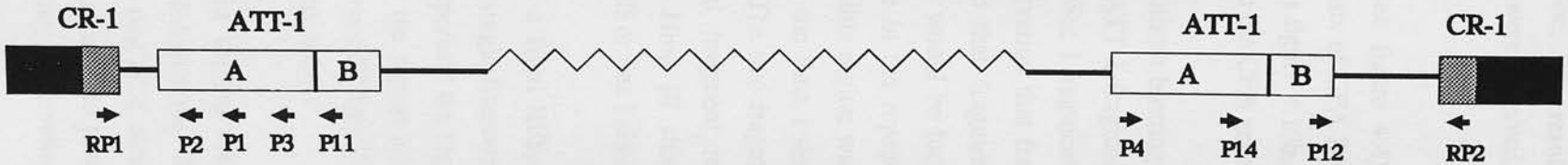
4.9 INVESTIGATION OF THE LOCATION OF ATT-1 USING RESTRICTION ENDONUCLEASES WHICH CUT IN THE REPEAT UNITS OF C500

4.9.1 Digestion using Sac I and Sac I/Hind III

CFA DNA was analysed by digestion with Sac I as the recognition sequence for this enzyme was present in the terminal repeat units of C500, hence the laddering

Figure 4.9

Diagramatic representation of the proposed CFA genomic arrangement



Key

- terminal repeats
- CR-1 repeat unit
- ATT-1

PCR products

5' RP1/P2 - 950bp
5' RP1/P1 - 1800bp
5' RP1/P3 - 2900bp
5' RP1/P11 - no products observed
3' RP2/P12 - 500bp
3' RP2/P14 - 1900bp
3' RP2/P4 - no products observed

associated with terminally located fragments did not occur. There was no recognition site for Sac I in the ATT-1 clone. The result of probing a Sac I CFA digest with ATT-1 was that two fragments were observed, one of 4.2kbp and another of 4.9kbp, as illustrated in figure 4.10a.

ATT-1 fragments A and B (see figure 4.4) were labelled with DIG then used individually to probe Sac I digests of CFA DNA. Both probes recognised the same Sac I fragments, as illustrated in figure 4.10b. This result implied that fragments A and B were each present twice in the CFA genome.

CFA DNA was analysed by Southern blotting after double digestion with Sac I and Hind III, and was probed with ATT-1 fragments A and B individually. Fragment B hybridised to two CFA Hind III/Sac I fragments, one of 0.9kbp and one of 1kbp (see figure 4.11, track 4). The observation that fragment B was present twice in a CFA Hind III/Sac I digest suggested that fragment B was present twice in the CFA genome. One of the fragments would be located between the Hind III site of the ATT-1 clone and the Sac I site in the repeat unit at the 3' terminus of the CFA genome. The predicted size of this region was approximately 1100bp (635bp from fragment B and 500bp between the Sma I site of ATT-1 and the Sac I site of the repeats, as predicted by PCR). The two fragments observed were both close to this predicted size. The additional fragment recognised by the probe presumably represented sequence from the Hind III site of the 5' terminal copy of ATT-1 downstream to the closest Hind III or Sac I site.

Fragment A was used to probe a Hind III/Sac I double digest of CFA DNA and hybridisation was observed to a single fragment of 3.1kbp (see figure 4.11, track 2). This fragment was believed to represent the region of DNA between the Hind III site of ATT-1 and the Sac I site of the repeat unit at the 5' end of the genome. The predicted size of this fragment was 3347bp (3197 bp of ATT-1 plus 150bp from the region between ATT-1 and the repeats).

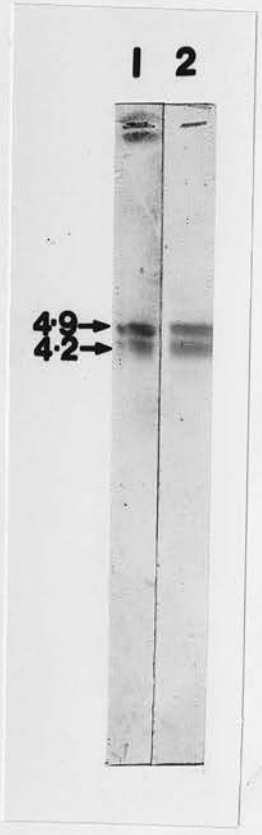
The PCR reactions described in section 4.8 had suggested that two copies of fragment A were present in the CFA genome. The single fragment observed in figure 4.11, track 2, was in contrast to this PCR data, and the two fragments exhibited in figure 4.10, when fragment A was used to probe a CFA Sac I digest. One possible explanation for these contradictory observations was that the 3.1kbp fragment in

Figure 4.10

Figure 4.10a
CFA digested with Sac I and probed with ATT-1

Track 1 - 1kb ladder
Track 2 - CFA digested with Sac I and probed with ATT-1

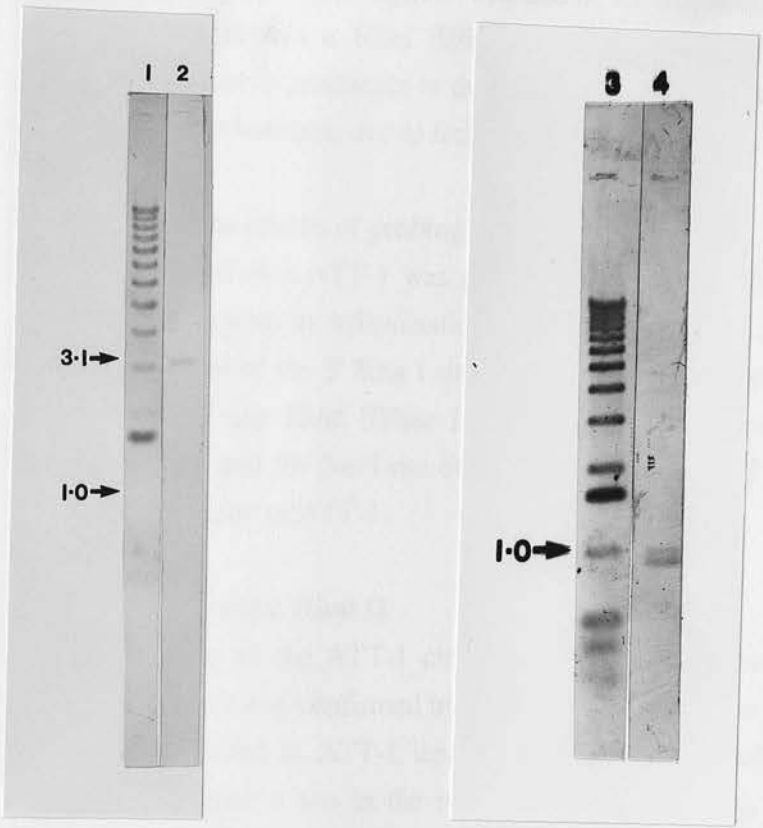
Figure 4.10b
CFA digested with Sac I and probed with ATT-1 fragments A and B



Track 1 - CFA digested with Sac I and probed with fragment A
Track 2 - CFA digested with Sac I and probed with fragment B

Figure 4.11

CFA digested with Hind III in conjunction with Sac I and probed with ATT-1 fragments A and B



Track 1 - 1kb ladder
Track 2 - CFA digested with Hind III/Sac I and probed with fragment A
Track 3 - 1kb ladder
Track 4 - CFA digested with Hind III/Sac I and probed with fragment B

figure 4.11 was a doublet, suggesting that a Sac I site is located 150bp 5' to the 3' terminal copy of ATT-1.

When ATT-1 was used to probe a Sac I digest of CFA DNA (see figure 4.10) fragments of 4.2 and 4.9kbp were observed. When the Hind III/Sac I fragments recognised by fragments A and B individually were added together the total sizes were 4.0 and 4.1kbp. This suggests that one of the fragments recognised by fragment A or fragment B was a Hind III/Hind III fragment rather than a Hind III/Sac I fragment. Further experiments to determine which fragment contained two Hind III sites were not undertaken, due to lack of suitable probe.

In conclusion, the results of probing Sac I CFA digests with ATT-1 and subclones of ATT-1 suggested that ATT-1 was present twice in the CFA genome. Analysis of Hind III/Sac I digests by hybridisation with the same probes indicated that the region directly upstream of the 5' Sma I site of ATT-1 was also repeated, since fragment A recognised only one Hind III/Sac I fragment. This implied that the 150bp region between ATT-1 and the Sac I site of the repeat units at the 5' end of the genome was repeated in addition to ATT-1.

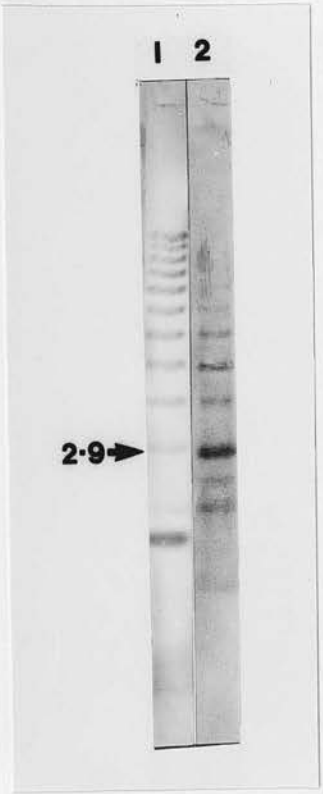
4.9.2 Digestion using Hind II

Mapping studies of the ATT-1 clone revealed a recognition site for Hind II, the presence of which was confirmed by sequencing (see chapter 3). The location of this Hind II site resulted in ATT-1 insert fragments of 975 and 2857bp. Hind II was known to recognise a site in the repeat unit of C500, with the result that Hind II digestion of CFA DNA had been observed to exhibit a supermolar fragment of approximately 1050bp representing the repeat unit (Bridgen *et al* 1989).

The ATT-1 insert, labelled with DIG, was used to probe a Hind II digest of CFA. A laddering effect was observed with a supermolar fragment of 2.9kbp evident, as illustrated in figure 4.12. This laddering effect was at first believed to be due to incomplete Hind II digestion of the CFA DNA. ATT-1 was hybridised to CFA digested with fresh aliquots of Hind II a further three times and on each occasion a ladder of fragments was observed. The consistent presence of Hind II ladders suggested that a proportion of the repeat units may be degenerate and fail to contain recognition sites for Hind II, although a proportion must contain Hind II sites to allow for the supermolar fragments observed on Hind II digestion. Variability of the repeat units of the AHV-1 isolate WC11 has previously been observed. Three

Figure 4.12

CFA digested with Hind II and probed with ATT-1



Track 1 - 1kb ladder
Track 2 - CFA digested with Hind II and probed with ATT-1

different sized components contribute to the terminal repeated region of this isolate (Bridgen *et al* 1989), therefore it is possible that C500 also contains repeat unit variability.

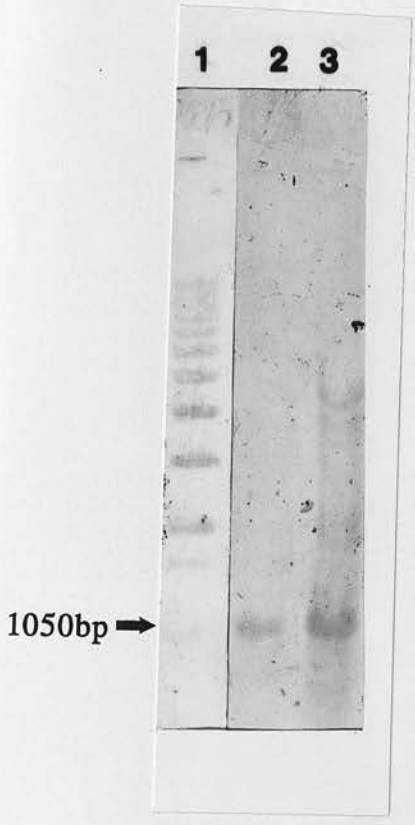
CFA DNA was digested with Sac I and Hind II respectively then probed with the CR-1 repeat clone. A single fragment of 1050bp, corresponding to the CR-1 clone, was observed in the Sac I track. In contrast a ladder of fragments was observed in the Hind II track, with the 1050bp fragment being of higher intensity (see figure 4.13). This further implied that the repeat units of the C500 genome were degenerate with respect to the presence of Hind II sites.

The ladder produced when ATT-1 was used to probe a Hind II CFA digest was, on closer examination, noted to be a double ladder, with one ladder being more intense than the other. This double laddering effect was similar to that observed when a Hind III digest of CFA was probed with ATT-1 (see figure 4.4a). The Hind II double laddering effect was investigated further by probing Hind II CFA digests with ATT-1 fragments A and B separately (see figure 4.14). An intense fragment of 2.9kbp was observed when fragment A was used as probe, in addition to the laddering effect. Fragment B hybridised to a fragment of approximately 1.9kbp and a lower intensity fragment of 2.4kbp, while additional larger fragments resembling a laddering effect were barely visible. The presence of a laddering effect in Hind II digests probed with ATT-1 fragments A and B respectively was consistent with the theory that ATT-1 is present twice in the CFA genome, located close to both termini.

The 5' terminal arrangement implied by PCR (see 4.8.3) resulted in a predicted terminal Hind II fragment of approximately 3.1kbp (comprised of 2857bp of ATT-1, 150bp between ATT-1 and CR1 and 24bp between the Hind II and Sac I sites of CR-1). This fragment was believed to be represented by the 2.9kbp supermolar fragment observed when fragment A was used to probe CFA DNA digested with Hind II. This 2.9kbp fragment was 200bp smaller than the predicted fragment, suggesting that a Hind II site was situated between the Sac I site of CR-1 and the Sma I site of ATT-1. The presence of a Hind II site was confirmed by Hind II digestion of the PCR product resulting from primer combination P2/RP1 (see 4.8.3). A fragment of less than 75bp was observed, corresponding to the 24bp between RP1 and the Hind II site in CR-1. In addition a fragment of 160bp was observed, implying that a Hind II site was present immediately upstream of the Sma I site of ATT-1. (A fragment of 740bp, representing the 5' ATT-1 DNA amplified in the PCR reaction was also observed).

Figure 4.13

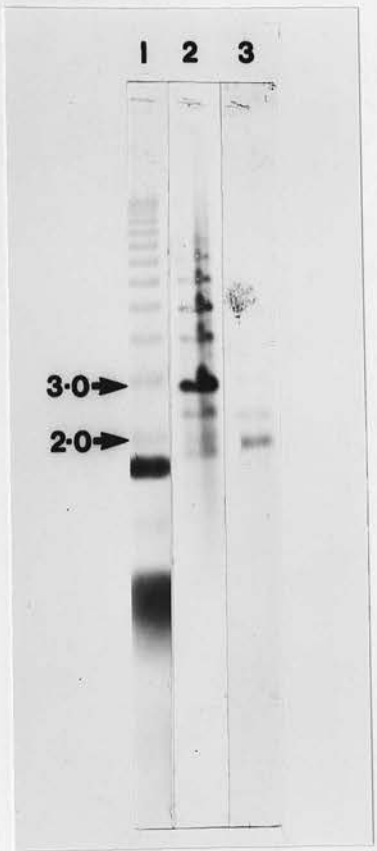
Hind II and Sac I digests probed with CR-1



Track 1 - 1kb ladder
Track 2 - CFA digested with Sac I and probed with CR-1
Track 3 - CFA digested with Hind II and probed with CR-1

Figure 4.14

CFA digested with Hind II and probed with ATT-1 fragments A and B



Track 1 - 1kb ladder
Track 2 - CFA digested with Hind II and probed with fragment A
Track 3 - CFA digested with Hind II and probed with fragment B

This result implied that the 2.9kbp supermolar fragment observed when fragment A was used to probe a Hind II digest of CFA DNA represented the 5' terminus of ATT-1.

The presence of a Hind II site between CR-1 and ATT-1 was difficult to interpret with respect to the laddering effect observed when fragment A was used as probe. If this Hind II site was the first Hind II site of the unique DNA at the 5' terminus of the genome then all ladders would be created by the sequence between this site and the repeats. Fragment A would not be expected to produce a laddering effect when used to probe a Hind II digest, as the Hind II site contained in fragment A would be the second site in the unique DNA rather than the first. The laddering effect observed when fragment A was used as probe was originally explained by the degeneracy of the repeats. Thus the Hind II site between CR-1 and ATT-1 may be part of the repeats hence also have a tendancy to degenerate. If this was the case then the ATT-1 clone must represent virtually the first base of the unique region of CFA C500 genome. The unique region may begin with small regions of copies of the repeat unit in a similar manner to that observed for HVS (Albrecht *et al* 1992), whereby a 116bp G+C rich region of DNA homologous to H-DNA lies between H-DNA and L-DNA. A similar arrangement to HVS would allow for second copies of Hind II and Sac I sites.

The observation that only one supermolar fragment was observed on probing a Hind II digest with fragment A was consistent with the single supermolar fragment observed on the equivalent Sac I/Hind III filter (see 4.9.1). Once again this could be interpreted such that fragment A was present as a single copy, or alternatively that the Hind II site immediately before ATT-1 was directly repeated at the 3' terminus of the genome.

Fragment B hybridised to fragments of 2.4 and 1.9kbp, once again illustrating that this fragment was present twice in the CFA genome. The 2.4kbp fragment was believed to represent the sequence between the Hind II site of ATT-1 and the Hind II site of the repeats at the 3' terminus of the genome. The predicted size of this fragment was 975bp from ATT-1, 500bp from the sequence between ATT-1 and RP2, and 950bp from RP2 to the Hind II site of CR-1, resulting in a total of approximately 2.4kbp. The submolar nature of this fragment was presumably due to the fact that the repeat units of CFA virus were degenerate, as previously discussed. The barely visible larger fragments represented the remainder of the 3' terminal

fragments containing fragment B. The 1.9kbp fragment was believed to represent the DNA downstream of the 5' location of ATT-1, such that a Hind II site was predicted 1kbp from the 3' Sma I site of ATT-1.

4.10 DISCUSSION

This chapter described experiments designed to locate the ATT-1 clone in the CFA genome. The ATT-1 clone was isolated from DNA which had undergone one or more alterations during passage in tissue culture. CFA virus represented the final *in vitro* arrangement of the C500 genome observed. As this derivative was apparently not altered on further serial passage it was possible to propagate larger amounts of the CFA derivative than its virulent counterparts. For this reason the location of ATT-1 was achieved prior to attempts to locate VIR-1 and VIR-2. The location of ATT-1 was intended to assist in the mapping studies of VIR-1 and VIR-2, by virtue of the 1923bp sequence it contained which was present in all three clones.

The location of ATT-1 was initially assessed by attempting to locate it using the WC11 map. This was unsuccessful therefore attempts were made to locate ATT-1 in the CFA genome. The location appeared to be terminal as when ATT-1 was used to probe CFA digested with various REs a laddering effect was observed. Further studies using Hind III demonstrated that either end of the ATT-1 clone produced a laddering effect, suggesting that the ATT-1 clone, or part of it, was located twice in the CFA genome.

The laddering patterns suggested various hypothetical arrangements of the ATT-1 clone in the CFA genome. The presence of the ATT-1 clone twice, at either end of the genome appeared the most likely arrangement. PCR was used to illustrate that ATT-1 was located close to the terminal repeats. The 1050bp terminal repeat unit was cloned and partially sequenced to generate oligonucleotides for use in PCR. The results of the PCR reactions described in this chapter demonstrated that the ATT-1 clone was located only 150bp from the Sac I site of the repeat unit at the 5' end of the unique region. Similarly the ATT-1 clone was located only 450bp from the Sac I site of the repeat unit at the 3' end of the unique region.

The use of PCR determined that sequences present in ATT-1 were located at either end of the genome. Attempts were made, using PCR, to prove that the entire ATT-1 clone was present at either end of the genome. Size constraints meant that PCR did

not demonstrate unequivocally that ATT-1 was present in its entirety at either end of the CFA genome. The PCR results generated did suggest that the entire ATT-1 clone was located at the 3' terminus, but did not prove fragment B to be adjacent to fragment A at the 5' terminus.

Further mapping studies used ATT-1 to probe CFA DNA digested with REs which recognised sites in the repeat unit of the molecule in order to prevent the laddering effect. The enzymes used were Sac I and Hind II.

The results of analysis of CFA DNA using Sac I and probing with the entire ATT-1 clone, and with fragments A and B of ATT-1, determined that the ATT-1 clone was present twice in the CFA genome.

Probing Hind II digests of CFA DNA with ATT-1, and subclones of ATT-1 demonstrated a laddering effect. A laddering effect should not have occurred with Hind II, as this enzyme recognised a site in the repeat unit of the C500 genome. The laddering effect was believed to be due to the degeneracy of the C500 repeat unit whereby not all the repeat units contained the Hind II site. The observation that a Hind II site was present between the 5' terminal repeats of the molecule and the Hind II site within the ATT-1 clone further complicated this reasoning. The Hind II site within this 150bp region was believed to be due to the presence of small regions of the repeat unit being present at the junction region between repeat and unique DNA. A possible explanation for the resultant laddering effect was that this region was also degenerate.

In conclusion, the experiments described in this chapter determined that the ATT-1 clone was present in the same orientation at either end of the genome.

Chapter 5
Investigations of the locations of VIR-1 and VIR-2

5.1 INTRODUCTION

This chapter describes molecular analysis of the C500 derivatives originally designated PP, CA and CFA. The aim of this analysis was to determine whether particular restriction endonucleases profiles were correlated with virulence and attenuation. The location of the ATT-1 clone in the CFA genome has been described in chapter 4. A similar strategy was employed to attempt to map the two clones derived from virulent C500 DNA, designated VIR-1 and VIR-2. The ATT-1 clone was previously used to probe restriction endonuclease digestions of virulent C500 DNA in order to compare the profiles with those observed when ATT-1 was used to probe equivalent CFA digests.

The three C500 derivatives were biologically cloned by Dr H.W. Reid. C500 DNA was probed with ATT-1, VIR-1 and VIR-2, both in the form of complete inserts and as subfragments. This work was limited by the low titre of virus obtained from tissue culture, resulting in yields of <5μg of viral DNA from harvests of low passage virus. In addition, the genomic alterations developing as a result of serial passage appeared very quickly during *in vitro* propagation (see chapter 3).

The genomic rearrangements occurring during serial passage were represented by the VIR-1, VIR-2 and ATT-1 molecular clones. These three clones all differed with regard to their 5' termini as illustrated in figure 3.11. This chapter investigates the significance of these non-homologous regions with regard to loss of virulence.

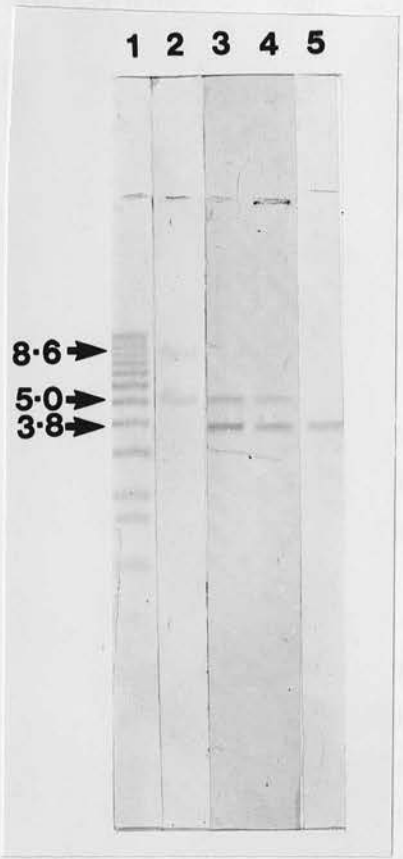
5.2 MOLECULAR ASSESSMENT OF VIRULENCE USING ATT-1 AS A PROBE

The virulence of the C500 populations designated PP, CA and CFA in chapter 3 was determined by using tissue culture harvests to inoculate rabbits as described in 2.6.3. Assessment of virulence was required at regular intervals during serial passage as the virus was known to become attenuated as a result of *in vitro* propagation.

The ATT-1 clone was used to probe Sma I digests of PP, CA, CF and CFA DNA respectively (see figure 5.1). Briefly, the ATT-1 clone recognised PP Sma I fragments of 8.6 and 5kbp, CA and CF Sma I fragments of 5 and 3.8kbp and a CFA Sma I fragment of 3.8kbp. The presence of the 5kbp fragment appeared to be related to virulence, as the only C500 derivative which failed to exhibit this fragment was

Figure 5.1

Sma I digests of C500 DNA probed with ATT-1



Track 1 - 1kb ladder
Track 2 - PP digested with Sma I
Track 3 - CA digested with Sma I
Track 4 - CF digested with Sma I
Track 5 - CFA digested with Sma I
Tracks 2-5 were probed with ATT-1 labelled with DIG

CFA, the attenuated derivative. In addition, this 5kbp fragment was also absent from the attenuated WC11 isolate (see figure 3.4).

5.3 DETERMINATION OF THE STATUS OF THE CA C500 DERIVATIVE

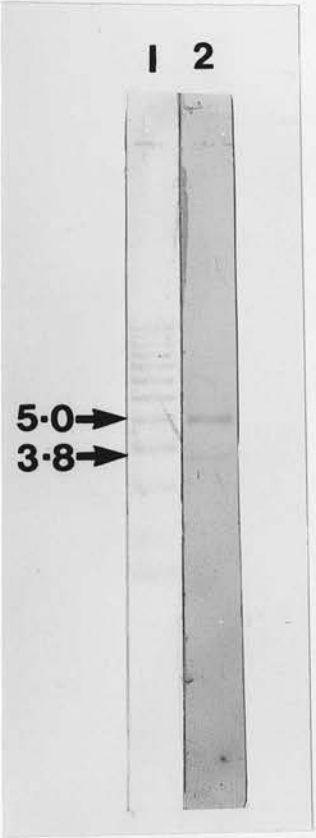
The CA population displayed fragments of 5 and 3.8kbp when digested with Sma I and probed with ATT-1. The presence of these two fragments suggested that the CA DNA represented either an intermediate genomic arrangement displaying characteristics noted in the PP and CFA derivatives or a mixed populution in which a proportion of the virions had retained virulence whilst the remaining virions had become attenuated. Closer examination of several preparations which had been designated CA revealed that the 5 and 3.8kbp fragments observed on ATT-1 hybridisation were not equimolar. Figure 5.2 illustrates one such CA derivative where the 5kbp fragment was supermolar to the 3.8kbp fragment. This lack of equimolarity suggested that the CA derivative was a mixed population containing two types of virion, one exhibiting some of the characteristics of the PP derivative, and the other exhibiting characteristics of the CFA derivative.

The nature of the CA derivative was investigated further by biological cloning. A CA preparation which exhibited the 5 and 3.8kbp fragments in an approximately equimolar ratio was biologically cloned by three serial limiting dilutions (kindly prepared by Dr H.W. Reid) to confirm that this variant was a mixed population rather than an intermediate genomic arrangement. Four biologically cloned derivatives were amplified in bovine turbinate cells to provide sufficient virus to allow genomic analysis. The DNA from each derivative was digested with Sma I then probed with VIR-1 (see figure 5.3). In three of the four samples fragments of 7 and 5kbp were observed. This suggested that the original CA preparation was a mixed population, and that the biological cloning resulted in only the virions displaying the 5kbp fragment (but not the 3.8kbp fragment) being recovered. The significance of the 7kbp fragment will be discussed in section 5.8.

Serial passage of the C500 isolate resulted in initially virulent cell-associated virus becoming attenuated cell-free virus. This alteration of the C500 population presumably occurred by selection of the derivatives most competent at replicating in the artificial environment of *in vitro* culture. This suggests that evolutionary pressure would select the attenuated virions from a mixed population because they are more replication competent than virulent virions. The biological cloning described in the

Figure 5.2

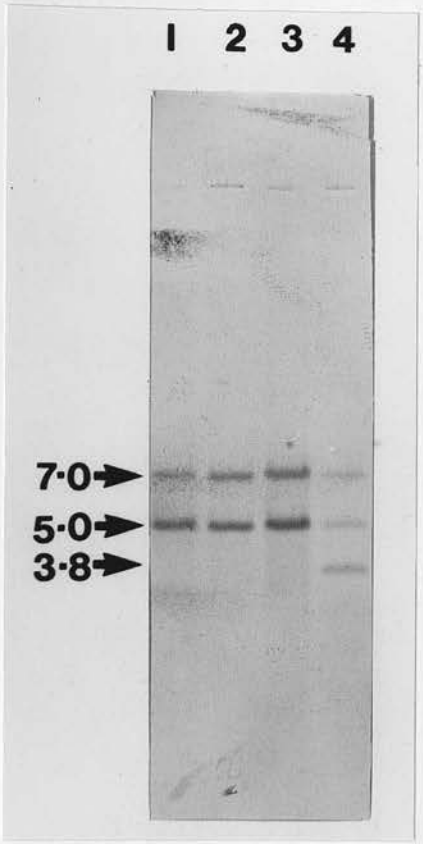
CA DNA digested with Sma I and probed with ATT-1



Track 1 - 1kb ladder
Track 2 - CA DNA digested with Sma I and probed with ATT-1

Figure 5.3

Biologically cloned CA derivatives probed with VIR-1



Track 1 - CA biologically cloned derivative 1 digested with Sma I
Track 2 - CA biologically cloned derivative 2 digested with Sma I
Track 3 - CA biologically cloned derivative 3 digested with Sma I
Track 4 - CA biologically cloned derivative 4 digested with Sma I
All tracks were probed with VIR-1

previous paragraph may have been expected to result in isolation of virions exhibiting CFA characteristics. The observation that three of the four biologically cloned CA derivatives did not exhibit the CFA 3.8kbp fragment was attributed to the low sample numbers used.

The fourth biologically cloned CA preparation exhibited fragments of 7, 5 and 3.8kbp on digestion with Sma I and subsequent probing with VIR-1. The presence of the 5kbp fragment suggested that, once again, it was the virions exhibiting some of the PP characteristics which had been selected by biological cloning. The presence of the 3.8kbp fragment was presumably due to a proportion of the virions undergoing the transition from virulence to attenuation in the amplification stage of viral preparation.

The CA derivative which was biologically cloned to produce the four DNA preparations described above was of a high passage level (37 passages in tissue culture). Previously derivatives of this passage level exhibited attenuation. The observation that this derivative had not become attenuated at this passage level was attributed to the initial biological cloning (to generate the PP derivative) which was undertaken immediately after isolation from a MCF affected rabbit, coupled with subsequent passage in small culture volumes. CF virus was detected in the supernatant of cell culture harvests of the four biologically cloned CA derivatives, resulting in recovery of an approximately equivalent quantity of DNA (0.1-1μg) to that recovered from the infected cells (i.e. CA DNA). One of the four derivatives was used to inoculate a rabbit and clinical signs of MCF were observed. Viral DNA extracted from a cell pellet derived from the lymph nodes of this animal exhibited a 5kbp fragment only when digested with Sma I and probed with ATT-1 (data not shown).

The experiments described above demonstrated that the CA derivative consisted of a mixed population of virions. In cell culture the presence of the PP derivative was transient, and the Sma I profile of the CA derivative was observed after very few passages. In general viral preparations failing to exhibit the 3.8kbp fragment were observed only at very low passage levels after isolation from affected rabbits and biological cloning. Limited amounts of DNA were prepared due to the small volume of material available from the cultures used to biologically clone the virus. As each amplification step required passaging the virus, the PP characteristic was quickly lost and the derivatives began to exhibit the 3.8kbp fragment in addition to the 7 and

5kbp fragments in the larger cell culture volumes required to prepare DNA in quantities of 5-10μg.

5.4 DETERMINATION OF THE STATUS OF THE PP DERIVATIVE

The PP derivative exhibited Sma I fragments of 8.6 and 5kbp when probed with ATT-1, as described in 5.2. This result was found to be inconsistent as the Sma I profiles of a proportion of biologically cloned PP derivatives exhibited only a 5kbp fragment when probed with ATT-1 (data not shown). The derivatives displaying the 8.6 and 5kbp Sma I fragments on probing with ATT-1 were subsequently referred to as PP1 derivatives, whilst those exhibiting the 5kbp fragment alone were referred to as PP2 derivatives. These results suggested that the PP2 derivative might represent an intermediate genomic arrangement, distinct from the PP1 and CFA derivatives. Alternatively, the population referred to as PP1 might be a mixed population, comprised of PP2 genomes and a second genomic arrangement.

The possibility that the PP1 derivative represented a mixed population was considered. If PP1 was a mixed population then this would imply that the ATT-1 clone only recognised sequence contained in a Sma I fragment of 8.6kbp in the earliest *in vitro* C500 population. The additional 5kbp fragment recognised in PP1 Sma I digests probed with ATT-1 would be due to PP2 virions. The PP1 derivatives had all been biologically cloned. This cloning should have isolated the population exhibiting the 8.6kbp fragment alone in a proportion of the independently derived PP derivatives if the 8.6kbp fragment was representative of an early genomic organisation. The 8.6 kbp fragment was never observed without the 5kbp fragment in any of the PP1 derivatives digested with Sma I and probed with ATT-1. This implied that if the 8.6kbp fragment was representative of an additional genomic organisation then the population it characterised must exist extremely transiently such that genomic rearrangements occurred almost immediately on extraction of virus from affected animals.

The absence of PP1 derivatives which demonstrated the 8.6kbp Sma I fragment alone suggested that PP1 represented a distinct population rather than a mixed population. In the PP1 genome ATT-1, or part of ATT-1, appeared to be present twice, once in the 8.6kbp fragment and again in the 5kbp fragment. This duplication of part of ATT-1 is consistent with the presence of ATT-1 twice in the CFA genome, as described in the previous chapter. Three of the four biologically cloned CA

derivatives described in section 5.3 exhibited fragments of 7 and 5kbp when digested with Sma I and probed with VIR-1. This population was distinct from the PP1 derivatives, which exhibited Sma I fragments of 8.6 and 5kbp when probed with VIR-1 (see figure 5.4), and therefore this population was designated PP2.

5.5 PROBING SMA I DIGESTS OF THE C500 DERIVATIVES WITH VIR-1 AND VIR-2

The hybridisation profiles described in this section, and the following three sections, are summarised in figure 5.22.

The VIR-1 clone was used to probe Sma I digests of PP, CA and CFA DNA (see figure 5.4). VIR-1 recognised the same fragments as ATT-1. In addition a fragment of 7kbp was recognised in the CA and CFA derivatives. Similarly, when the VIR-1 clone was used to probe a Sma I digest of PP2 a fragment of 7kbp was observed in addition to a 5kbp fragment (see figure 5.3). The observation that the 7kbp PP2, CA and CFA fragment was present with VIR-1 but not with ATT-1 suggested that this fragment was recognised by the region of the VIR-1 clone which was not present in ATT-1 (i.e. the 5' end of the VIR-1 clone). The significance of the 7kbp Sma I fragment will be discussed in section 5.8. The Sma I fragment profiles using VIR-1 as probe were found to be consistent for a number of independently derived PP1, PP2, CA and CFA derivatives.

Comparison of the ATT-1 and VIR-1 Sma I C500 hybridisation profiles also suggested that the CA population represented both the PP2 and CFA derivatives. The PP2 derivative became a mixed population of PP2 and CFA virions very rapidly, therefore it was not possible to prepare quantities of PP2 DNA in excess of 2-5µg. In the remainder of this study a proportion of the PP2 characteristics were determined from CA preparations by eliminating the fragments attributed to CFA virions.

The VIR-2 clone was used to probe Sma I digests of PP1, CA and CFA derivatives (see figure 5.5). These hybridisations demonstrated that the VIR-2 clone produced an identical pattern to that observed with the ATT-1 clone (see figure 5.1). The 7kbp fragment recognised by VIR-1 was not observed, implying that the 7kbp Sma I fragment was recognised by the region of VIR-1 which did not demonstrate homology to VIR-2.

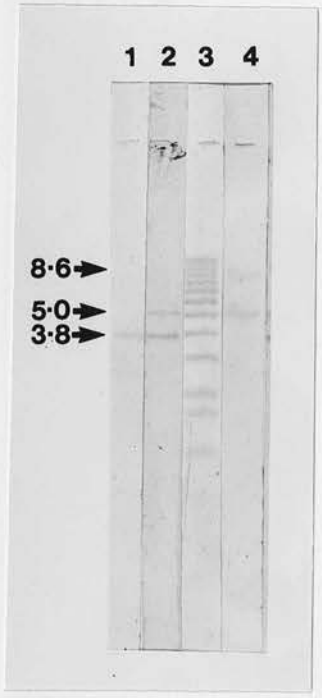
Figure 5.4

Sma I profiles of AHV-1 digests probed with VIR-1

Track kb -1kb ladder
Track 1 -WC11 digested with Sma I
Track 2 -CFA digested with Sma I
Track 3 -CA digested with Sma I
Track 4 -PP digested with Sma I
Tracks 2-4 were probed with VIR-1

Figure 5.5

Sma I digests of C500 DNA probed with VIR-2



Track 1 - CFA DNA digested with Sma I
Track 2 - CA DNA digested with Sma I
Track 3 - 1kb ladder
Track 4 - PP DNA digested with Sma I
Tracks 1, 2 and 4 were probed with VIR-2

5.6 PROBING HIND III DIGESTS OF THE C500 DERIVATIVES WITH ATT-1

When ATT-1 was used to probe Hind III digests of PP1 and PP2 a faint laddering effect was observed. In addition supermolar fragments of 3.6kbp were observed in both derivatives. The ATT-1 clone had produced a double ladder when used to probe a Hind III digest of CFA DNA (see figure 4.4a). ATT-1 fragments A and B also each created a laddering effect when used to probe CFA Hind III digests (see figures 4.4b and 4.4c). These fragments were used individually to probe Hind III digests of PP1 DNA in order to compare the Hind III hybridisation profiles of the PP1 and CFA genomes.

When fragment A was used to probe a Hind III digest of PP1 DNA a doublet of 3.6kbp was observed. This was in contrast to the laddering effect observed when a CFA Hind III digest was probed with fragment A (see figure 5.6). This 3.6kbp doublet was believed to represent the VIR-1 and VIR-2 clones, which differ by only 10bp and both contain sequence homologous to ATT-1, suggesting that VIR-1 and VIR-2 were both present in the PP1 genome.

A single fragment of 3.6kbp was observed when fragment A was used to probe a Hind III digest of PP2 DNA (data not shown), suggesting that either VIR-1 and VIR-2, but not both, were present in this population.

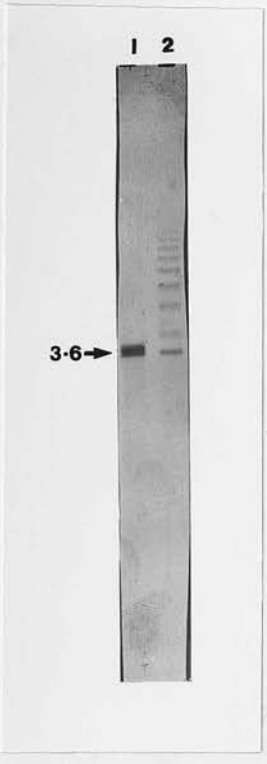
The absence of a laddering effect in both of the PP derivatives suggested the VIR-1 and VIR-2 clones were not located as close to the 5' terminal repeats of the PP1 and PP2 genomes as fragment A was to the 5' terminal repeats of the CFA genome.

Hind III digests of PP1 and PP2 were probed with ATT-1 fragment B. This fragment was present in ATT-1 but not VIR-1 or VIR-2. A laddering effect was observed in both PP derivatives (data not shown), suggesting that fragment B was present and was terminally located in both PP1 and PP2.

The possible terminal location of VIR-1 and/or VIR-2 in the PP1 genome was investigated further by probing Bam HI and Eco RI PP1 digests with VIR-1. The VIR-1 and VIR-2 sequences do not contain recognition sites for either of these enzymes. A laddering effect was observed in both digests, as illustrated in figure 5.7. This suggested that either VIR-1 or VIR-2 was terminally located, or that both clones

Figure 5.6

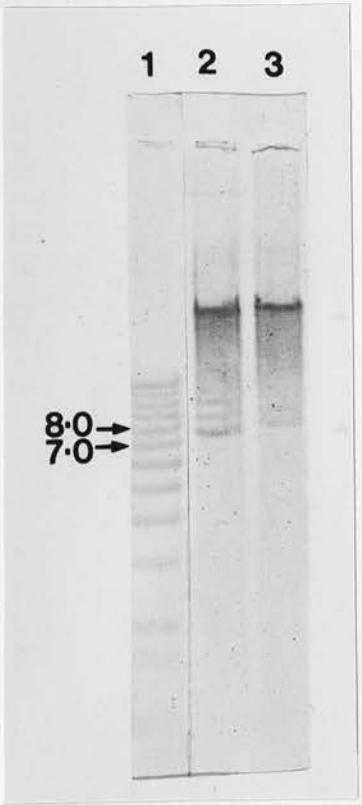
Hind III C500 digests probed with ATT-1



Track 1 - PP1 digested with Hind III
Track 2 - CFA digested with Hind III
Both tracks were probed with fragment A

Figure 5.7

PP1 DNA digests probed with VIR-1



Track 1 - 1kb ladder
Track 2 - PP1 digested with Bam HI
Track 3 - PP1 digested with Eco RI
Tracks 2 and 3 were probed with VIR-1

were terminally located at opposite ends of the PP1 genome. The smallest Bam HI fragment observed was approximately 7.5kbp, and the smallest Eco RI fragment was approximately 8kbp. These fragment sizes suggested that a Bam HI site was located a maximum of 5kbp from the start of the terminal repeats at one end of the genome (assuming that at least one repeat unit of 1050bp was required for each virion), and that an Eco RI site was located a maximum of 7kbp from the start of the terminal repeats.

5.7 PROBING HIND III DIGESTS OF THE C500 DERIVATIVES WITH VIR-1

The VIR-1 clone was used to probe PP1, CA and CFA Hind III digests as illustrated in figure 5.8. In PP1 DNA a fragment of 3.6kbp was observed, whilst in CA and CFA DNA laddering occurred. The smallest fragment evident in the CA and CFA digests was 2kbp. The laddering effect in the CA and CFA digests was due to the location of fragment A at the 5' terminus of the CFA genome, since the CFA genome was present in the CA population. The VIR-1 clone recognised ATT-1 fragment A, as the two clones shared homology for the 1923bp of fragment A. This region of homology is illustrated in figure 5.9, and will subsequently be referred to as fragment C. When the VIR-1 clone was used to probe a PP2 digest (see figure 5.10) fragments of 3.6 and 2kbp were observed.

The 2kbp fragment observed in the PP2, CA and CFA digests was not observed when ATT-1 was used to probe equivalent C500 digests. The CR-1 insert, representing the repeat unit of C500, also did not recognise this 2kbp fragment (see figure 5.10). This implied that the 2kbp Hind III fragment was not the smallest fragment of the ladder due to the repeats, but was recognised specifically by the VIR-1 clone. Hybridisation to this 2kbp Hind III fragment was attributed to all or part of the 5' 1616bp region of the VIR-1 clone, since this is the region of VIR-1 which is not homologous to ATT-1 and VIR-2 (i.e. see figure 5.9).

The presence of the 2kbp fragment in PP2, CA and CFA Hind III digests when probed with VIR-1 suggested that the 5' 1616bp region of VIR-1 may be present twice in these C500 derivatives. One copy of this sequence would be present in the arrangement represented by the VIR-1 clone, and the second copy would be involved in the 2kbp Hind III fragment. Alternatively the conformation of VIR-1 might be absent in PP2, CA and CFA being replaced by the 2kbp fragment.

were terminally located at opposite ends of the PP1 genome. The smallest Bam HI fragment observed was approximately 7.5kbp, and the smallest Eco RI fragment was approximately 8kbp. These fragment sizes suggested that a Bam HI site was located a maximum of 5kbp from the start of the terminal repeats at one end of the genome (assuming that at least one repeat unit of 1050bp was required for each virion), and that an Eco RI site was located a maximum of 7kbp from the start of the terminal repeats.

5.7 PROBING HIND III DIGESTS OF THE C500 DERIVATIVES WITH VIR-1

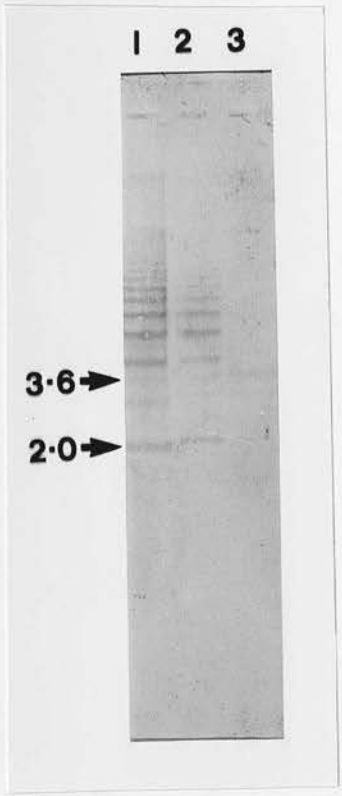
The VIR-1 clone was used to probe PP1, CA and CFA Hind III digests as illustrated in figure 5.8. In PP1 DNA a fragment of 3.6kbp was observed, whilst in CA and CFA DNA laddering occurred. The smallest fragment evident in the CA and CFA digests was 2kbp. The laddering effect in the CA and CFA digests was due to the location of fragment A at the 5' terminus of the CFA genome, since the CFA genome was present in the CA population. The VIR-1 clone recognised ATT-1 fragment A, as the two clones shared homology for the 1923bp of fragment A. This region of homology is illustrated in figure 5.9, and will subsequently be referred to as fragment C. When the VIR-1 clone was used to probe a PP2 digest (see figure 5.10) fragments of 3.6 and 2kbp were observed.

The 2kbp fragment observed in the PP2, CA and CFA digests was not observed when ATT-1 was used to probe equivalent C500 digests. The CR-1 insert, representing the repeat unit of C500, also did not recognise this 2kbp fragment (see figure 5.10). This implied that the 2kbp Hind III fragment was not the smallest fragment of the ladder due to the repeats, but was recognised specifically by the VIR-1 clone. Hybridisation to this 2kbp Hind III fragment was attributed to all or part of the 5' 1616bp region of the VIR-1 clone, since this is the region of VIR-1 which is not homologous to ATT-1 and VIR-2 (i.e. see figure 5.9).

The presence of the 2kbp fragment in PP2, CA and CFA Hind III digests when probed with VIR-1 suggested that the 5' 1616bp region of VIR-1 may be present twice in these C500 derivatives. One copy of this sequence would be present in the arrangement represented by the VIR-1 clone, and the second copy would be involved in the 2kbp Hind III fragment. Alternatively the conformation of VIR-1 might be absent in PP2, CA and CFA being replaced by the 2kbp fragment.

Figure 5.8

C500 Hind III digests probed with VIR-1



Track 1 - CFA digested with Hind III
Track 2 - CA digested with Hind III
Track 3 - PP1 digested with Hind III
All tracks were probed with VIR-1

Figure 5.9

Diagramatic representation of VIR-1, VIR-2 and ATT-1

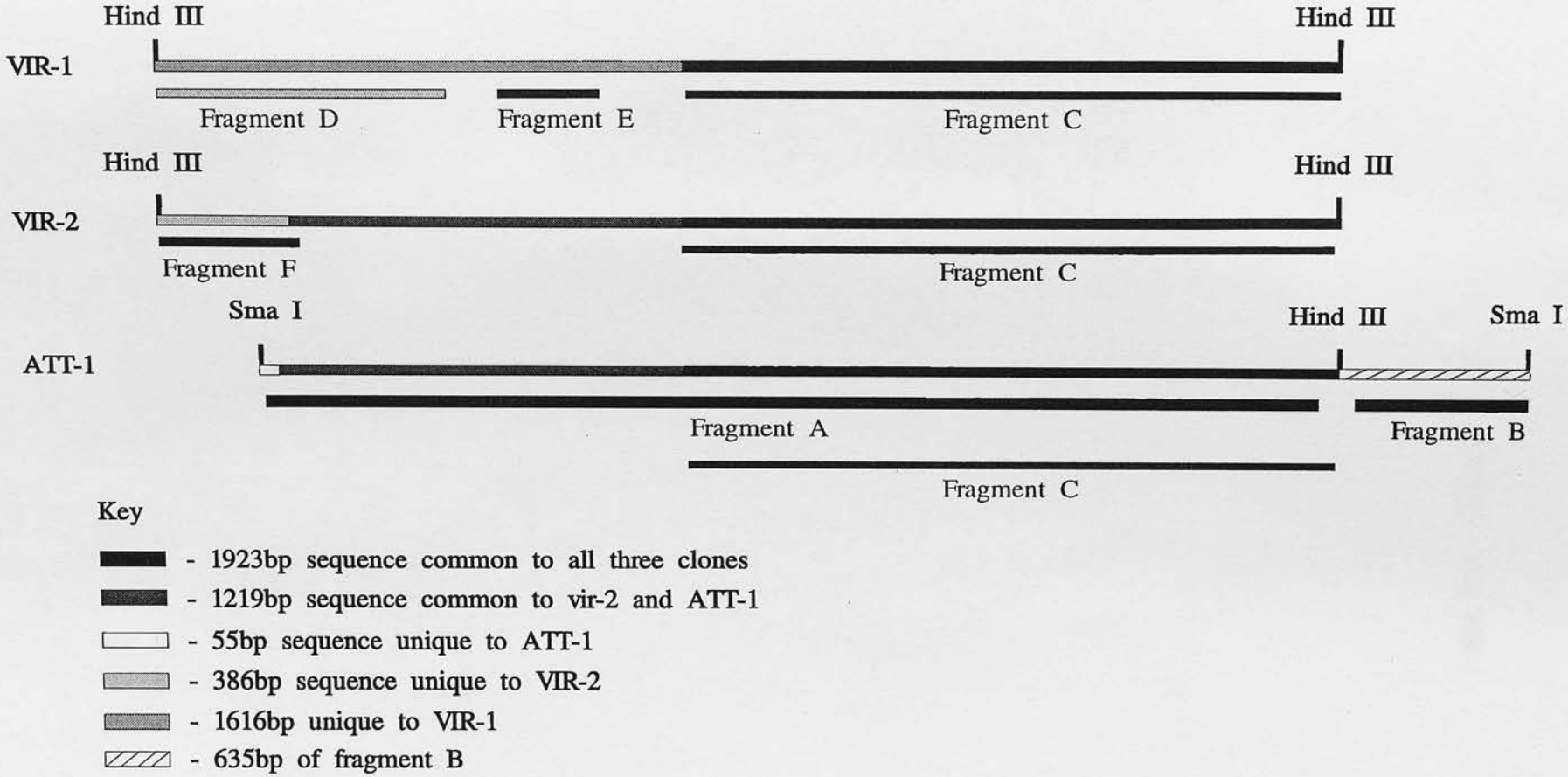
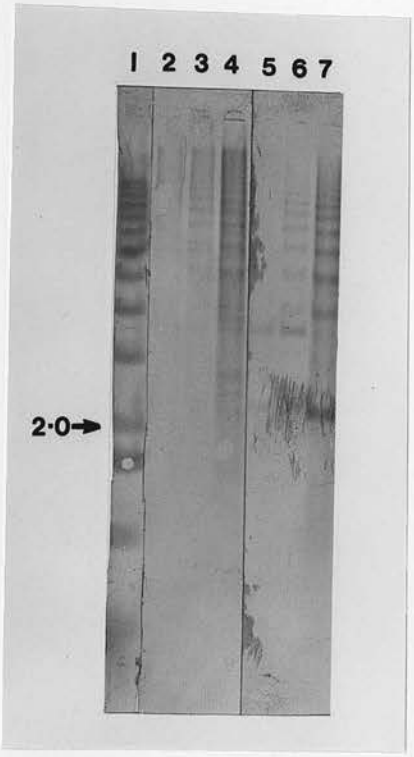


Figure 5.10

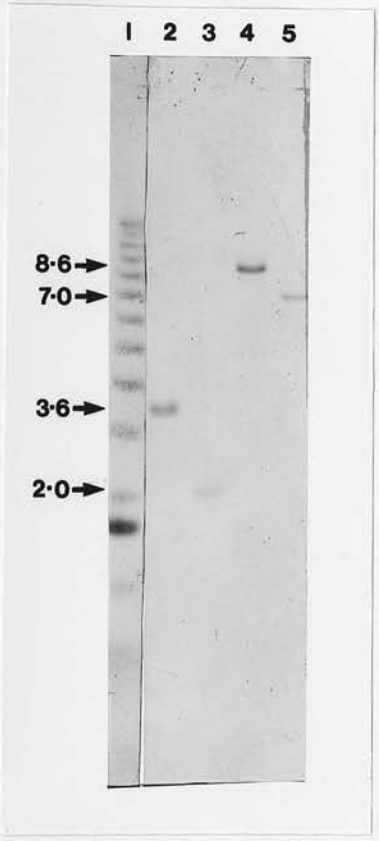
Hind III digests of C500 DNA probed with CR-1 and VIR-1



Track 1 - 1kb ladder
Track 2 - PP2 digested with Hind III
Track 3 - CA digested with Hind III
Track 4 - CFA digested with Hind III
Track 5 - PP2 digested with Hind III
Track 6 - CA digested with Hind III
Track 7 - CFA digested with Hind III
Tracks 2-4 were probed with CR-1
Tracks 5-7 were probed with VIR-1

Figure 5.11

C500 RE digests probed with fragment D



Track 1 - 1kb ladder
Track 2 - PP1 digested with Hind III
Track 3 - CFA digested with Hind III
Track 4 - PP1 digested with Sma I
Track 5 - CFA digested with Sma I
Tracks 2-5 were probed with VIR-1 fragment D

was recognised specifically by the "unique" region of VIR-1. The observation that ATT-1 and VIR-2 did not recognise this 7kbp fragment implied that fragment C (i.e. the sequence common to all three clones) was not contained within this 7kbp Sma I fragment. The observation that fragment D recognised only one CFA Sma I fragment (7kbp) and that this fragment did not contain fragment C, suggested that the CFA genome did not contain the VIR-1 clone conformation. Similarly a fragment of 7kbp was observed when VIR-1 was used to probe PP2 and CA Sma I digests, but was not evident when ATT-1 and VIR-2 were used to probe equivalent digests (compare figures 5.1, 5.4 and 5.5). This confirmed the Hind III data that the VIR-1 clone conformation was absent from the PP2 and CA genomes.

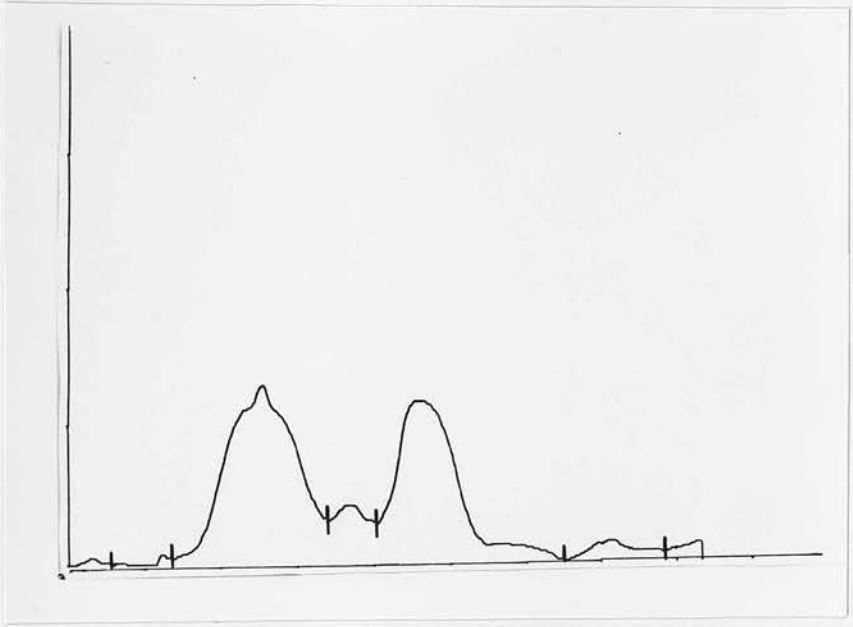
An additional VIR-1 "unique" probe, fragment E in figure 5.9, was used to probe further aliquots of the PP1 and CFA Hind III and Sma I digests. An identical hybridisation profile to that produced using fragment D was observed, suggesting that the majority of the 1616bp VIR-1 "unique" sequence was present in the CFA genome.

Analysis of the Sma I hybridisation profiles using the various probes described in this chapter resulted in preliminary conclusions regarding the genomic rearrangements which occur in the C500 genome as a result of *in vitro* propagation. The sequence common to all three clones (i.e. fragment C in figure 5.9) appeared to be present twice in the PP1 genome, contained in both the 8.6 and 5kbp Sma I fragments. Since the sequence "unique" to VIR-1 recognised the 8.6kbp Sma I fragment, coupled with the assumption that VIR-1 and VIR-2 were both present in PP1 DNA (represented by the doublet in track 1, figure 5.6), it was postulated that the sequence of VIR-1 was encoded by the Sma I 8.6kbp fragment, and the sequence of VIR-2 was contained within the 5kbp fragment.

Analysis of a Sma I digest of PP1 DNA which had been probed with ATT-1 (track 2, figure 5.1) using a Joyce Loebl densitometer on reflectance mode resulted in two peaks (see figure 5.12). The ATT-1 clone shared homology with VIR-1 over a region of 1923bp (fragment C). In contrast, ATT-1 shared homology with VIR-2 over a region of 3134bp. The ratio of the two Sma I peaks when probed with ATT-1 was therefore predicted to be 1923:3134, or 1:1.63. The observed results were that the 5kbp peak was 1.7 times as intense as the 8.6kbp peak. This implied that VIR-2 was contained in the 5kbp peak, and that VIR-1 was contained in the 8.6kbp peak.

Figure 5.12

Densitometry plot of PP1 digested with Sma I and probed with ATT-1



The first peak corresponds to the 5kbp Sma I fragment
The second peak corresponds to the 8.6kbp Sma I fragment

The hybridisation profile of PP2 on Sma I digestion exhibited a 7kbp fragment, but no 8.6kbp fragment, when probed with VIR-1. This 7kbp fragment was not observed when ATT-1 and VIR-2 respectively were used to probe PP2 DNA digested with Sma I, suggesting that the copy of fragment C contained in VIR-1 was lost as PP1 derivatives developed characteristics of PP2 derivatives following serial passage. However, fragment C was present twice in the CFA genome, in the form of the two copies of ATT-1 present in this derivative (see chapter 4). The apparent loss, then subsequent re-emergence of the "second" copy of fragment C during *in vitro* propagation of C500 virus will be discussed in section 5.12.

5.9 THE USE OF PCR TO ATTEMPT TO LOCATE VIR-1 AND VIR-2 IN THE PP GENOMES

5.9.1 Introduction

PCR was used to aid in the location studies described in chapter 4. Similar PCR reactions were carried with the PP1 and PP2 genomes as target DNA, using the oligonucleotides designed for sequencing VIR-1, VIR-2 and ATT-1 in conjunction with the CR-1 repeat clone primers, RP1 and RP2. The primers used and the clones from which they were derived are illustrated diagramatically in figure 5.13. The 5' and 3' termini of the PP1 and PP2 genome are arbitrarily labelled in the same orientation as the CFA genomes.

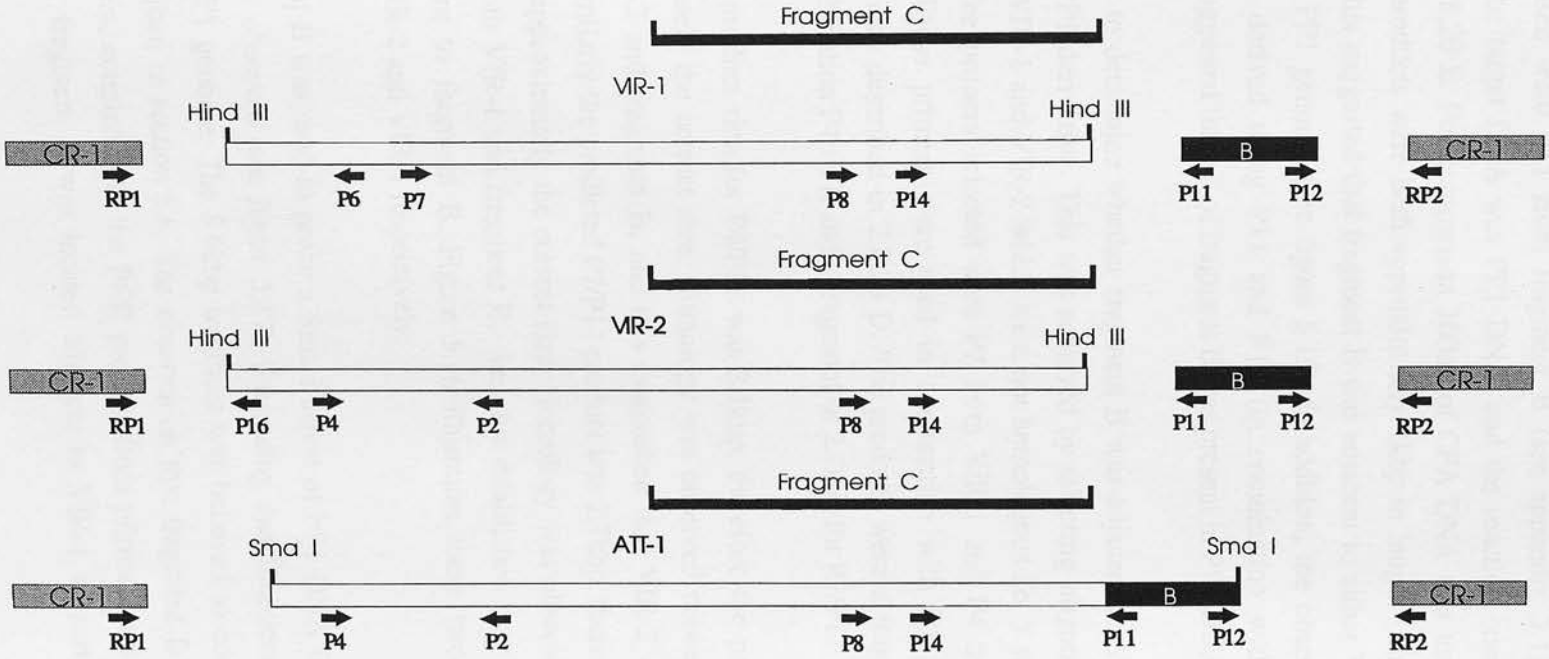
5.9.2 The use of PCR to investigate whether VIR-1 or VIR-2 was adjacent to the 3' terminus of PP1.

When ATT-1 fragment B was used to probe a PP1 Hind III digest a laddering effect was observed. This suggested that the PP1 equivalent to ATT-1 fragment B was terminally located in the PP1 genome. The experiments described in chapter 4 illustrated that ATT-1 fragment B was located adjacent to the arbitrarily determined 3' terminus of the CFA genome. PCR was used in order to determine whether VIR-1, VIR-2, or both were adjacent to fragment B, and whether fragment B was terminally located, in a similar manner to the arrangement of ATT-1 at the 3' end of the CFA genome.

A primer from fragment B, P12, (see appendix 3 and figure 5.13) was used in conjunction with RP2 following conditions described in 2.8.20 A, with PP1 DNA as target. The resultant product was 500bp in length, identical to the CFA product using these primers. The PP1 and CFA P12/RP2 products were probed with CR-1 and

Figure 5.13

Diagramatic representation of the C500 cloned fragments and the primers used for PCR



Fragment c represents the region common to all three clones
This figure illustrates the primers used in PCR to investigate the possibility of terminal location

homology was observed (see figure 5.14). This suggested that the PP1 fragment B equivalent was terminally located (see figure 5.15).

The location of VIR-1 and VIR-2 in relation to fragment B was assessed by PCR. The primers used were P11 from fragment B (see appendix 3.1) and P8, from fragment C. The target DNA was PP1 DNA, and the reaction conditions were as described in 2.8.20 E. For comparison 100ng of CFA DNA was treated identically The resultant products were both approximately 1kbp in length (illustrated later in figure 6.10). This suggested that fragment B was adjacent to either VIR-1 or VIR-2 or both in the PP1 genome (see figure 5.15). In addition, the observation that the products were derived using P11 and P12 (in combination with P8 and RP2 respectively) suggested that all of fragment B was present in PP1 DNA.

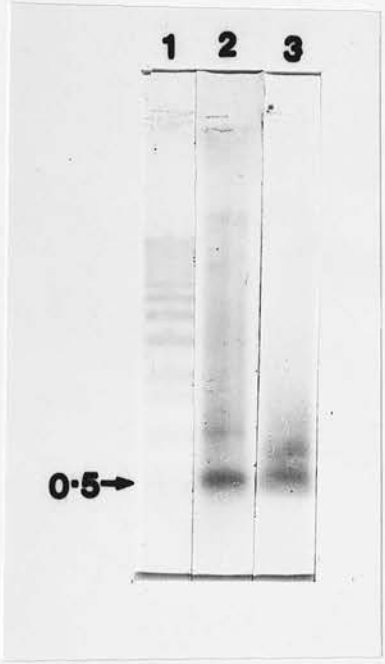
PCR was used to determine whether fragment B was adjacent to VIR-1, VIR-2 or both in the PP1 derivative. This was achieved by selecting oligonucleotides from the regions of VIR-1 and VIR-2 which were not homologous i.e. 5' to fragment C in both clones. The primers selected were P7 from VIR-1 and P4 from VIR-2 (see figure 5.13). These primers were used in conjunction with P11, following the reaction conditions described in 2.8.20 D. The products were a fragment of 3.1kbp for primer combination P4/P11, and a fragment of 2.7kbp for P7/P11.

The predicted product size for P4/P11 was 3.1kbp, therefore the product observed was approximately the correct size. Homology was observed between this product and both VIR-2 and fragment B, and this established that VIR-2 was adjacent to fragment B. Similarly the predicted P7/P11 product was 2.7kbp, therefore the product observed was approximately the correct size. Homology was observed between this product and both VIR-1 and fragment B, and this established that VIR-1 was also located adjacent to fragment B. Figure 5.16 illustrates these two PCR products probed with VIR-2 and VIR-1 respectively.

When fragment B was used to probe a Sma I digest of PP1 DNA fragments of 8.6 and 5kbp were observed (see figure 5.17), illustrating that fragment B was present twice in the PP1 genome. The 8.6kbp fragment was believed to contain the VIR-1 clone as described in section 5.8. The observation that fragment B recognised this 8.6kbp fragment, coupled with the PCR product from primer combination P7/P11, suggested that fragment B was located adjacent to VIR-1 as part of this 8.6kbp

Figure 5.14

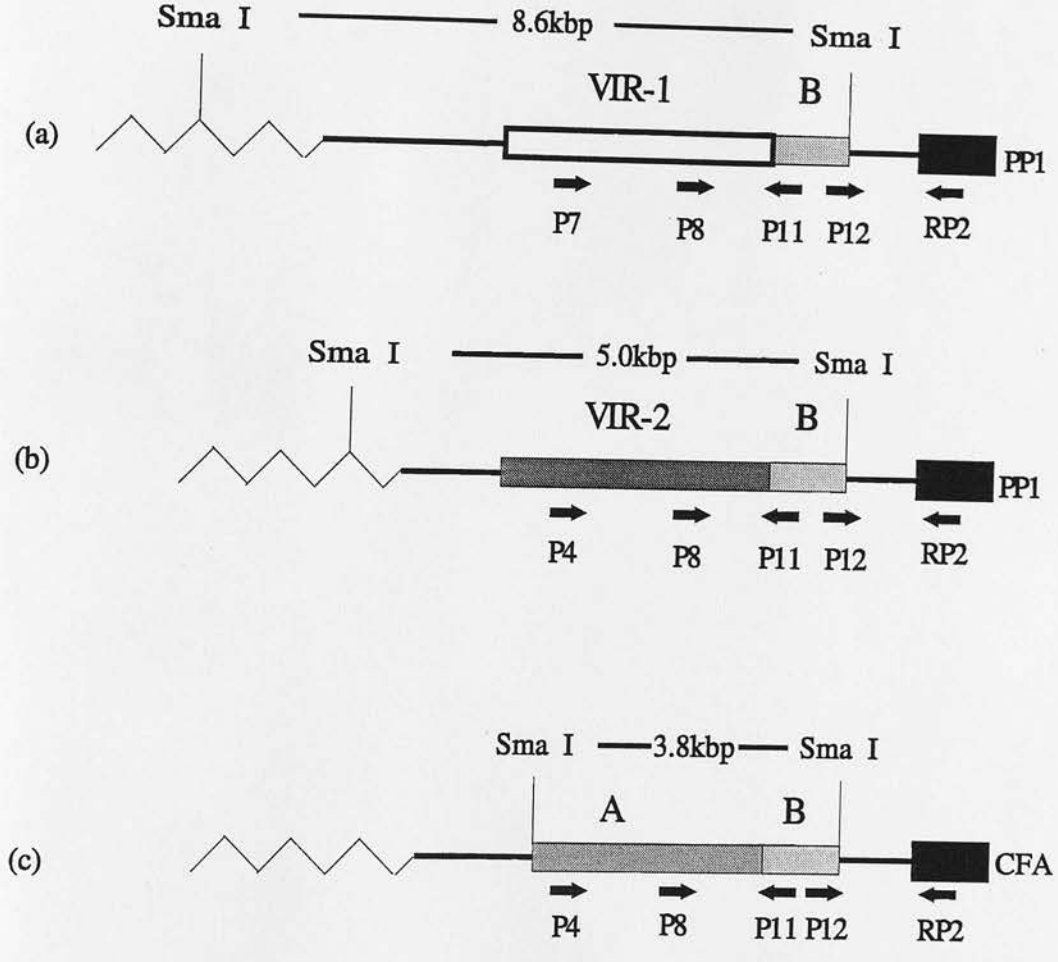
Comparison of 3' terminal PP1 and CFA PCR reactions



Track 1 - 1kb ladder
Track 2 - PCR product of PP1 amplified with primer combination P12/RP2
Track 3 - PCR product of CFA amplified with primer combination P12/RP2
Tracks 2 and 3 were probed with CR-1 labelled with DIG

Figure 5.15

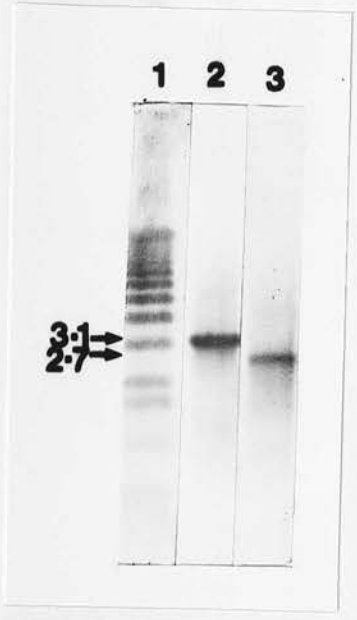
Possible 3' termini locations of PP1



This figure illustrates the possible 3' termini of the PP1 genome in (a) and (b), and compares this arrangement to that established for CFA in chapter 4

Figure 5.16

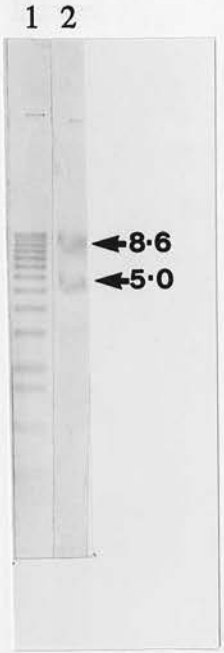
PP1 PCR products



Track 1 - 1kb ladder
Track 2 - PCR product of PP1 amplified using primer combination P4/P11
Track 3 - PCR product of PP1 amplified using primer combination P7/P11
Track 1 was probed with VIR-2
Track 2 was probed with VIR-1

Figure 5.17

PP1 DNA digested with Sma I and probed with fragment B



Track 1 - 1kb ladder
Track 2 - PP1 digested with Sma I and probed with fragment B

fragment. Similarly, the 5kbp fragment apparently contained the VIR-2 clone, located adjacent to fragment B.

The PCR reaction involving P12 and RP2 suggested that fragment B was terminally located. Figure 5.15 illustrates the two possible 3' genomic arrangements of the PP1 derivative from the data described in this section, compared to the CFA 3' terminus which was established in chapter 4. PCR reactions were carried out using RP2 with a primer common to VIR-1 and VIR-2, P8 (see figure 5.13), with PP1 DNA as target. The reaction conditions were as described in 2.8.20 F, and a product of approximately 2kbp was observed. This product was homologous to ATT-1 fragment B (data not shown), further suggesting that either VIR-1 or VIR-2 was located close to the 3' terminal repeats of the PP1 genome.

5.9.3 The use of PCR to investigate whether part of VIR-1 or VIR-2 was adjacent to the 3' terminus of PP2

The VIR-1 clone conformation was apparently absent from the PP2 genome, but part or all of the 5' 1616bp of the sequence "unique" to VIR-1 was present. If PP2 was a direct descendant of PP1, and if the genomic arrangement illustrated in figure 5.15a was correct then at least 1923bp (fragment C) were absent from the 3' terminal location in PP2 compared to PP1. If this loss was a straight deletion then gene amplification using P7 in combination with P11 would result in a product of <2.7kbp (with 2.7kbp being the PP1 product). A PCR reaction involving P7/P11 was carried out following the conditions described in 2.8.20 A, with 100ng of PP2 DNA as the target sequence, and a product of 500bp was observed. Hybridisation was not observed between this product and fragment D, suggesting that the 500bp fragment may be an artefact of PCR. In order to investigate this further P7 and P11 were used alone in two separate PCR reactions, following the conditions described for P7 in conjunction with P11. The results of these reactions were that a 500bp fragment was observed in the sample using P11 alone, implying further that the P7/P11 product was an artefact.

The lack of products observed using primer combination P7/P11 with PP2 DNA suggested that fragment B may not be adjacent to the region of VIR-1 which was included in the PP2 genome (i.e. the 2kbp Hind III fragment). The PP1 derivative contained two copies of fragment B, as illustrated in figure 5.17. In contrast the PP2 derivative appeared to contain only one copy of fragment B, since ATT-1 (which contains fragment B) recognised only the Sma I 5kbp fragment in PP2 DNA. This

suggested that it was not only a copy of fragment C that was absent from PP2, when compared to PP1, but also a copy of fragment B. This possibility was investigated further by using P7 in combination with RP2 in a PCR reaction with 100ng of PP2 as target DNA, following the reaction conditions described in section 2.7.20 D. A 500bp product was observed, but this product did not exhibit homology to either CR-1 or VIR-1, therefore was assumed to be an artefact of PCR.

A PCR reaction involving P12 and RP2, following the reaction conditions described for CFA P12/RP2 in 2.8.20 A, was used to amplify 100ng PP2 DNA. The product observed was identical in size to the fragments illustrated in figure 5.14, and exhibited homology to CR-1 (data not shown). In addition, a PCR reaction involving P8 and RP2, following the reaction conditions described in 2.8.20 F, was used to amplify PP2 DNA. A product homologous to ATT-1 fragment B, and of approximately 2kbp in length, was observed (data not shown). This implied that VIR-2 was located directly upstream to the 5' terminal location of fragment B, since the region of VIR-1 encoding P8 was absent from the PP2 genome.

A PCR reaction involving primer combination P4/P11 following the conditions described in 2.8.20 D with PP2 DNA as target resulted in a product of 3.1kbp. This product was identical in size to the PP1 equivalent product (see figure 5.16), and exhibited homology to VIR-2 (data not shown). This was predicted because the VIR-2 clone and fragment B were located within the PP2 5kbp fragment, and this fragment was apparently unchanged between the PP1 and PP2 derivatives.

5.9.4 The use of PCR to investigate whether VIR-2 was adjacent to the 5' termini of PP1 and PP2

The VIR-2 and ATT-1 clones differed in the sequence located at their 5' termini such that the VIR-2 clone contained a 386bp region of DNA which was not present in ATT-1. When ATT-1 and VIR-2 were aligned (see figure 5.9) this 386bp fragment was located to the 5' side of ATT-1. If the VIR-2 clone was located adjacent to the 5' terminus of PP1 and PP2 this 386bp region, together with the Sma I data, implied that VIR-2 could not be located as close to the 5' termini of the PP1 and PP2 genomes as ATT-1 was to the 5' terminus of the CFA genome.

The PCR reaction which had confirmed that ATT-1 fragment A was terminally located at the 5' end of the CFA genome involved primers P2 and RP1. The PP1 and PP2 derivatives respectively were amplified using these oligonucleotides, following

the reaction conditions described in section 2.8.20 A. No products were observed in either reaction, suggesting that VIR-2 was located further from the 5' termini of both PP derivatives than ATT-1 was from the CFA 5' terminus.

The lack of product using P2 and RP1 to amplify PP1 DNA contrasted with an earlier result when P2 and RP1 were used in an identical reaction with virulent CA DNA (see figure 4.7). This PCR using 100ng CA DNA as input generated a 950bp product which was identical in size to the equivalent CFA product. This product was thought to be due to the presence of CFA virions in the CA population, as described in section 5.3.

A primer closer to the 5' terminus of VIR-2, P16 (see appendix and figure 5.13), was used in combination with RP1, as P16 was closer to the 5' terminus of VIR-2, although not present in ATT-1. The reaction conditions were as described in 2.8.20 A, using 100ng of PP1 and PP2 respectively as target DNA. A 400bp product was observed, but this product did not exhibit homology to either VIR-2 or CR-1, therefore was apparently an artefact of PCR.

5.9.5 The use of PCR to investigate whether VIR-1 (or the 5' region of VIR-1 present in PP2) was adjacent to the 5' termini of PP1 and PP2

The VIR-1 clone did not contain the sequence for P2, nor did its 2kbp Hind III PP2 equivalent, therefore an alternative 5' VIR-1 primer, P6 (see figure 5.13) was used in conjunction with RP1 in an identical reaction cycle to that described for RP1 and P2. No products were observed, suggesting that VIR-1 and its PP2 equivalent were located further from the 5' terminus of both PP derivatives than ATT-1 was from the CFA 5' terminus.

5.9.6 The use of PCR to investigate whether VIR-1 or VIR-2 was adjacent to the 5' termini of PP1 and PP2 in an inverted orientation

The possibility that the VIR-1 or the VIR-2 clone might be adjacent to the 5' terminal repeats of PP1 and PP2 in an inverted orientation to that of ATT-1 at the 5' terminus of the CFA genome was considered. An inverted orientation would explain the lack of product observed in sections 5.9.4 and 5.9.5. A potential theory for the rearrangement involving a "flipping" mechanism would be possible if VIR-1 or VIR-2 was orientated in this manner. When P14 was combined with RP1, using PP1 DNA as target, and following the reaction conditions described in 2.8.20 A, no

products were observed. This result was inconclusive as the fragments may have been larger than the size constraints.

5.9.7 Conclusions

PCR was used to assess whether the locations of VIR-1 and VIR-2 in the PP1 and PP2 genomes paralleled the location of ATT-1 in the CFA genome. The conclusions from these PCRs were that ATT-1 fragment B was present twice in the PP1 genome, adjacent to both VIR-1 and VIR-2, and was present at least once in the PP2 genome, adjacent to VIR-2. In addition fragment B was shown to be adjacent to the 3' termini of both virulent genomes. In the PP1 genome either VIR-1 or VIR-2 was located 5' to the 3' terminal location of fragment B, and in the PP2 genome, VIR-2 was located 5' to the 3' terminal location of fragment B.

The PCR reactions to determine whether VIR-1 or VIR-2 were located at the 5' termini of both PP derivatives, in the orientation presented in figure 5.13, or in the opposite orientation, did not result in products. The lack of products did not demonstrate conclusively that VIR-1 and VIR-2 were not located close to the 5' terminal repeats of PP1 and PP2, but did illustrate that neither clone was located as close to the 5' repeats of either PP derivative as ATT-1 was to the 5' terminus of the CFA genome.

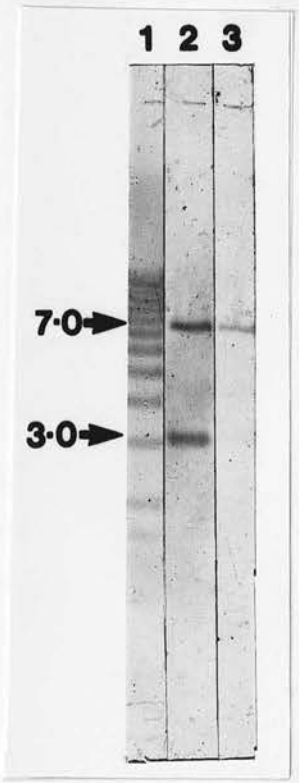
5.10 SAC I ANALYSIS OF PP1

PP1 was digested with Sac I and probed with VIR-1 and VIR-2 respectively, as illustrated in figure 5.18. VIR-1 recognised fragments of 7.5 and 3kbp, while VIR-2 recognised a single fragment of 7.5kbp. These results were difficult to interpret as VIR-1 was predicted to hybridise to three fragments, one corresponding to VIR-2 and two corresponding to VIR-1 since VIR-1 has internal Sac I sites (the central 167bp VIR-1 Sac I fragment was not observed). Similarly VIR-2 was predicted to hybridise to two Sac I fragments, one corresponding to itself, and one corresponding to the homologous region of VIR-1 (fragment C in figure 5.9). The observed result suggested that VIR-1 and VIR-2 must be located very close to each other.

The observation that VIR-2 recognised only one fragment (of 7.5kbp) in a PP1 Sac I digest suggested that either the VIR-1 and VIR-2 copies of fragment C were contained in identically sized Sac I fragments, or that both copies were contained in the one Sac I fragment. The data described in the previous section suggested that both

Figure 5.18

PP1 DNA digested with Sac I and probed with VIR-1 and VIR-2



Track 1 - 1kb ladder
Track 2 - PP1 digested with Sac I and probed with VIR-1
Track 3 - PP1 digested with Sac I and probed with VIR-2

VIR-1 and VIR-2 were located adjacent to fragment B in PP1. In addition, the Sma I hybridisation profiles suggested that VIR-2 was included in a Sma I 5kbp fragment, and that VIR-1 was included in a Sma I 8.6kbp fragment. One possible organisation of the above data is presented in figure 5.19. The PP1 Hind III, Sac I and Sma I data is combined to generate the sizes of each individual fragment. The Sma I 5 and 8.6kbp fragments are illustrated below this map, and the Sac I fragments are illustrated above it. The observed Sac I fragment recognised by both VIR-1 and VIR-2 was 7.5kbp, however, inaccuracies do occur when sizing fragments from the results of Southern blots, therefore this fragment may be closer to the 8.1kbp fragment calculated.

The PP1 genomic organisation proposed in figure 5.19 was investigated further using PCR with various primers 5' and 3' to the 800bp fragment located between VIR-2 and the copy of fragment B adjacent to VIR-1. Although extensive PCR analysis involving the VIR-2 primers P16 and P2, in combination with the fragment B primer 12, and VIR-1 primer 14 was undertaken, product was not generated. One possible explanation for this failure was that the genome is not arranged in the manner described in figure 5.19. An alternative explanation is that the nature of the sequence encoded by this 800bp region somehow inhibits the PCR, possibly by its secondary structure as has been observed previously (Green and Sargan 1991). This region of the C500 genome also proved impossible to clone (see section 3.6).

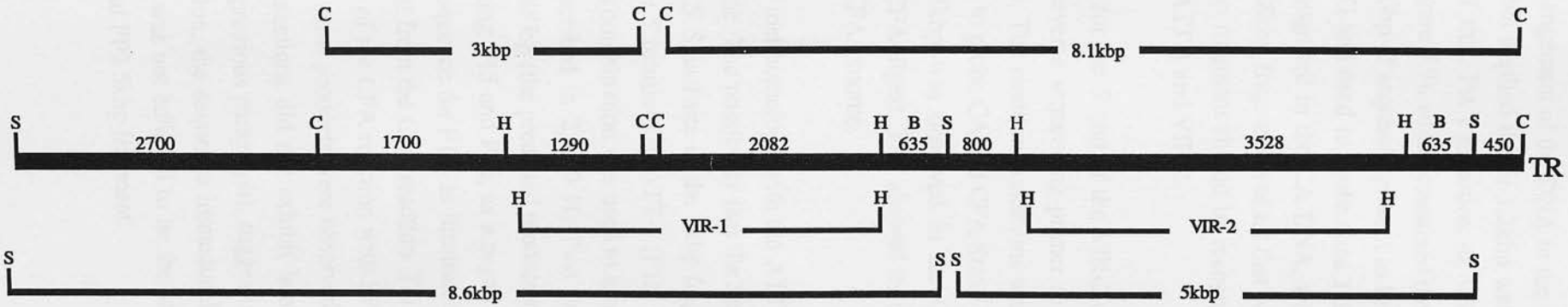
5.11 PROBING C500 DIGESTS WITH THE REGION OF VIR-2 ABSENT FROM VIR-1 AND ATT-1

The 5' 386bp of the VIR-2 clone did not share homology with the ATT-1 or VIR-1 clones. This fragment is referred to as the sequence "unique" to VIR-2 in the remainder of this section.

The VIR-2 sequence was believed to be present in the 5kbp Sma I fragment observed in PP1 and PP2 DNA (see section 5.8). Earlier results had suggested that this fragment was correlated with virulence (see section 5.2). The 5kbp Sma I fragment was absent from the CFA genome, and was apparently replaced by a 3.8kbp Sma I fragment. No other fragment was observed in the CFA digest when VIR-2 or ATT-1 were used as probe, suggesting that at least 1.2kbp of DNA which was present in the PP1 and PP2 derivatives was absent from CFA genome.

Figure 5.19

The proposed arrangement of VIR-1 and VIR-2 in the PP1 genome



Key
C - Sac I
H - Hind III
S - Sma I
TR- terminal repeats
B - fragment B from ATT-1

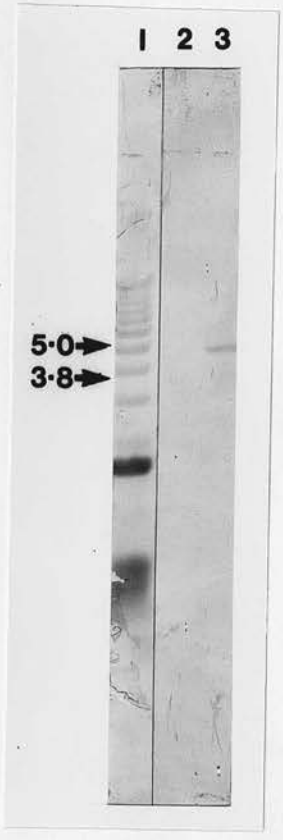
The proposed arrangement of the DNA in the 5kbp Sma I fragment was as illustrated in figure 5.19. This implied that if 1.2kbp was "lost" then the VIR-2 "unique" DNA would be part of this. Pst I digestion of VIR-2 produced a 5' fragment of 496bp (fragment F in figure 5.9), which contained the 386bp of sequence "unique" to VIR-2 in addition to 110bp of sequence present in both VIR-2 and ATT-1. Fragment F was labelled with DIG and used to probe Sma I digests of CA and CFA DNA. A 5kbp fragment was recognised in the CA DNA, with a faint 3.8kbp fragment (see figure 5.20a). A faint 3.8kbp fragment was all that was observed in the CFA track. In both digests the 3.8kbp fragment should be recognised by the 110bp of the probe which was common to ATT-1 and VIR-2.

A probe specific for the 5' end of the VIR-2 clone was prepared using P16 and the pBS SK+ M13 reverse sequencing primer to generate a DIG labelled probe using PCR (see 2.8.21). The reaction conditions were as described in 2.8.20 G. The PCR product was used to probe CA and CFA Sma I digests respectively (see figure 5.20b). A fragment of 5kbp was observed in the CA digest, while no fragments were observed in the CFA digest. This showed that the 5' region of the VIR-2 clone was absent from the CFA genome.

The VIR-2 clone lost homology with the ATT-1 clone 55bp before the 5' Sma I site of the ATT-1 clone. The possibility that the Sma I site and the 55bp 3' to it, might be conserved as the 5' Sma I site of the 5kbp fragment was investigated. A primer was designed from the 5' terminus of ATT-1 (P15) and used in conjunction with P16 (see figure 5.21). This combination was used to amplify 100ng of PP1 DNA. The reaction cycle was as described in 2.8.10 H. Two products were observed from the PP1 reaction, one of 800bp (the predicted product size) and one of 200bp. CFA DNA was also amplified using P15 and P16, as a negative control, since the CFA genome did not contain the sequence for P16, as illustrated above. However, a product of 800bp was also observed from the CFA reaction. This result was initially thought to be due to contamination of the CFA reaction with PP1 DNA, therefore both reactions were repeated, but the same products were observed. The observed products from the PP1 and CFA PCR reactions did not exhibit homology to a VIR-2 specific probe (as described in the previous paragraph), suggesting that this product was an artefact of PCR. In conclusion, the sequence immediately 3' to the 5' Sma I site of the CFA 3.8kbp fragment was not believed to be the same as the sequence 3' to the 5' Sma I site of the PP1 and PP2 5kbp fragment.

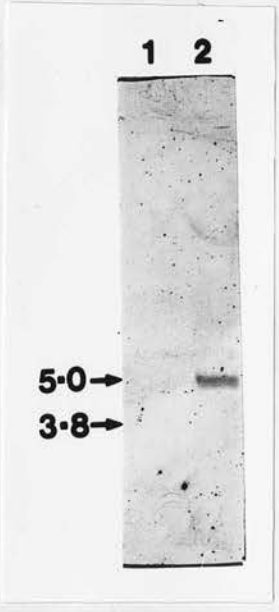
Figure 5.20
Sma I C500 digests probed with 5' VIR-2 sequence

Figure 5.20a
Probing CA and CFA Sma I digests with fragment F



Track 1 - 1kb ladder
Track 2 - Sma I digested CFA
Track 3 - Sma I digested CA
Tracks 2 and 3 were probed with fragment F

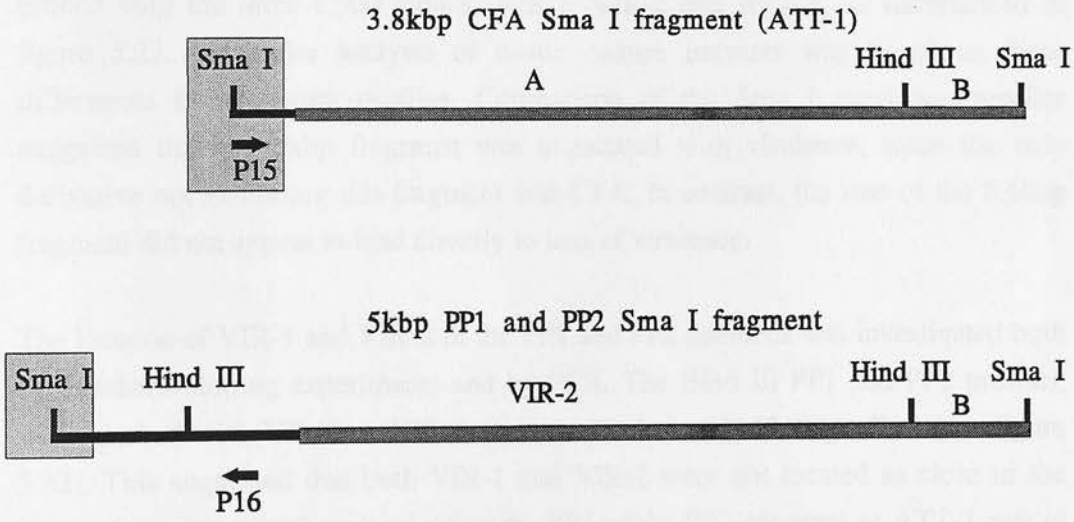
Figure 5.20b
Probing CA and CFA Sma I digests with P16 VIR-2 fragment



Track 1 - Sma I digested CFA
Track 2 - Sma I digested CA
Tracks 1 and 2 were probed with VIR-2 P16 5' terminal fragment

Figure 5.21

Diagramatic representation of the potentially conserved region 3' to the 5' Sma I site in the CFA 3.8kbp Sma I fragment and the PP1/PP2 5kbp Sma I fragment



5.12 DISCUSSION

This chapter described attempts to establish the location of the VIR-1 and VIR-2 clones in virulent C500 viral genomes. In chapter three the virulent derivatives of C500 were described as PP, CA and CF (see figure 3.2). Closer investigation of the population designated PP illustrated that two separate genomic arrangements were present, therefore the PP derivative was divided into PP1 and PP2 genomes. Similarly, closer examination of the CA population showed it to consist of PP2 and CFA genomes. The CF population was not discussed specifically in this chapter, although the limited data presented suggested that this population was similar to the CA population (see figure 5.1, where the CA and CF profiles are identical).

Restriction digests, using Hind III and Sma I respectively, resulted in consistent, but distinct restriction profiles for the three C500 derivatives, PP1, PP2 and CFA, when probed with the three C500 clones, VIR-1, VIR-2 and ATT-1, as summarised in figure 5.22. Molecular analysis of tissue culture harvests was based on these differences in restriction profiles. Comparison of the Sma I restriction profiles suggested that the 5kbp fragment was associated with virulence, since the only derivative not exhibiting this fragment was CFA. In contrast, the loss of the 8.6kbp fragment did not appear to lead directly to loss of virulence.

The location of VIR-1 and VIR-2 in the PP1 and PP2 genomes was investigated both by Southern blotting experiments and by PCR. The Hind III PP1 and PP2 profiles, when probed with VIR-1 or VIR-2, did not result in a laddering effect (see figure 5.22). This suggested that both VIR-1 and VIR-2 were not located as close to the terminal repeats at both ends of either the PP1 or the PP2 genomes as ATT-1 was to both terminal repeats in the CFA genome.

The use of PCR to investigate the location of VIR-1 and VIR-2 established a 3' terminal location for either VIR-1 or VIR-2, adjacent to fragment B in the PP1 genome, and a 3' location for VIR-2 adjacent to fragment B in the PP2 genome. The PCR reactions to locate either VIR-1 or VIR-2 at the 5' end of the PP1 and PP2 genomes were inconclusive. However, amalgamation of PCR and Southern blotting data determined that neither VIR-1 nor VIR-2 was located as close to the 5' terminal repeats of the PP1 and PP2 genomes as ATT-1 was to the 5' terminal repeats of the CFA genome.

Figure 5.22
Summary of restriction profiles of the C500 derivatives

PROBE	DERIVATIVE	Sma I	Hind III
VIR-1	PP1	8.6 and 5kbp	3.6kbp
VIR-2	PP1	8.6 and 5kbp	3.6kbp
ATT-1	PP1	8.6 and 5kbp	3.6kbp + laddering
VIR-1	PP2	7 and 5kbp	3.6 and 2kbp
VIR-2	PP2	5kbp	3.6kbp
ATT-1	PP2	5kbp	3.6kbp + laddering
VIR-1	CFA	7 and 3.8kbp	2kbp + laddering
VIR-2	CFA	3.8kbp	laddering
ATT-1	CFA	3.8kbp	laddering

The hybridisation profiles observed when VIR-1 and VIR-2 were used to probe a Sac I digest of PP1 DNA presented the possibility that, rather than being located at opposite ends of the genome, VIR-1 and VIR-2 were located almost adjacent to each other (see figure 5.19).

The PP2 genome was shown to differ from the PP1 genome in that PP2 DNA appeared to contain only one copy of the 1923bp sequence common to all three clones (i.e. fragment C), while PP1 contained two copies of this sequence, as VIR-1 and VIR-2. In addition, PP2 appeared to contain only one copy of the 3' 635bp Hind III/Sma I ATT-1 fragment (i.e. fragment B). This suggested that a minimum of approximately 2.6kb present in PP1 was absent from PP2. Comparison of the PP1 and PP2 Sma I profiles, when probed with VIR-1, suggested a loss of 1.6kbp, as the PP1 Sma I 8.6kbp fragment was replaced by a 7kbp Sma I fragment. Similarly the VIR-1 3.6kbp Hind III PP1 fragment appeared to be replaced by a 2kbp PP2 fragment. Although these Sma I and Hind III fragments represent a rearranged region of the C500 genome recognised by the 5' sequence of VIR-1, direct comparison of fragment sizes does not necessarily reflect loss of sequence.

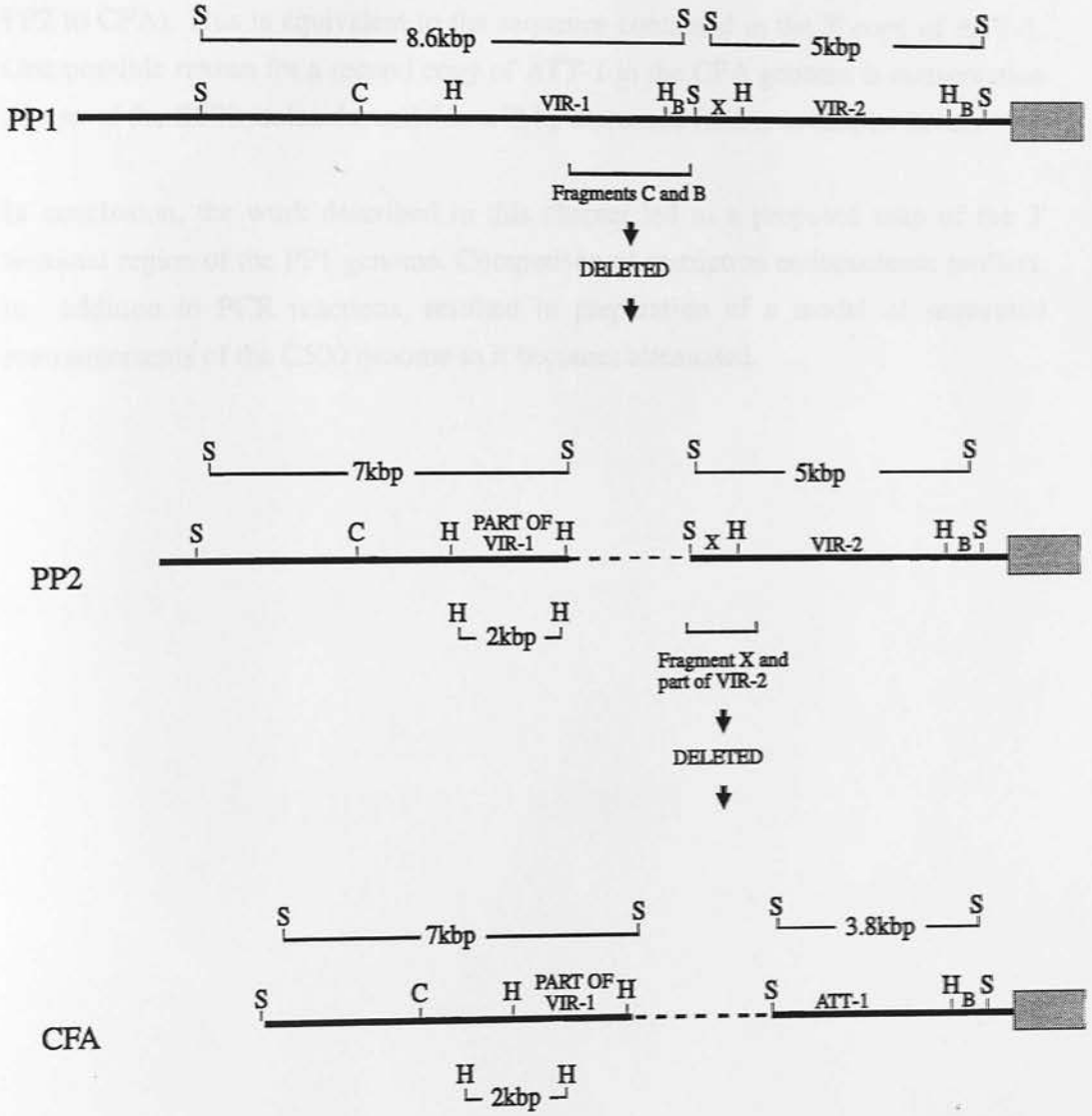
The proposed 3' terminal organisation of the PP1 genome was used to create a potential organisation for the 3' terminal sequence of the PP2 genome (see figure 5.23), taking into consideration the PP2 Sma I 7kbp and the Hind III 2kbp fragments observed. This organisation was not assessed by PCR due to lack of PP2 DNA. The arrangement presented is only one of several possiblities.

Comparison of the PP2 and CFA Sma I profiles illustrated that a 5kbp PP2 fragment became a 3.8kbp CFA fragment, suggesting a region of >1.2kbp was absent from the CFA genome. The VIR-2 clone was shown to contain 386bp (at the 5' end) which was present in PP2, but absent in CFA DNA. This suggested that a deletion of >386bp had occurred. The PCR reaction to determine whether the ATT-1 5' Sma I site was conserved as the 5' Sma I site of the PP2 5kbp fragment suggested that it was not, therefore the sequence deleted was probably >1.2kbp. One possible arrangement of the 3' terminus of the CFA genome is presented in figure 5.23.

While substantial work will be required to confirm the suggested genomic organisations presented in figure 5.23, they do present a possible model of two sequential deletion events occurring in the transition from virulence to attenuation of the C500 isolate of AHV-1.

Figure 5.23

Proposed rearrangements of the C500 genome



Key
C - Sac I
H - Hind III
S - Sma I
--- rearranged DNA
▒ - terminal repeats

The CFA genome contains a second copy of the ATT-1 clone at the 5' end of the genome. Neither this copy of ATT-1, nor an equivalent, appear to be located at the 5' end of the PP1 or PP2 genomes. The total DNA lost sequentially from the PP1 to the CFA genomic arrangement is >3.8kbp (i.e.2.6kbp from PP1 to PP2, and 1.2kbp from PP2 to CFA). This is equivalent to the sequence contained in the 5' copy of ATT-1. One possible reason for a second copy of ATT-1 in the CFA genome is conservation of size of the C500 molecule, and this will be discussed further in chapter seven.

In conclusion, the work described in this chapter led to a proposed map of the 3' terminal region of the PP1 genome. Comparison of restriction endonuclease profiles, in addition to PCR reactions, resulted in preparation of a model of sequential rearrangements of the C500 genome as it becomes attenuated.

Chapter 6
Sequence Analysis

6.1 INTRODUCTION

The three C500 clones, ATT-1, VIR-1 and VIR-2 were all sequenced (see chapter 3), and the complete sequences are presented in appendix 3. Sequence analysis was performed using several of the computer programmes which form part of the University of Wisconsin genetics computer group (GCG) sequence analysis package.

This chapter describes the various programmes which were used to analyse the C500 sequence. The sequence generated was compared to the sequences of other herpesviruses, particularly the gammaherpesviruses *Herpesvirus saimiri* (HVS) and Epstein-Barr virus (EBV), both of which have been completely sequenced (Albrecht *et al* 1992, Baer *et al* 1984). Comparison to other related rhadinoviruses, such as OHV-2, was not possible as the OHV-2 genome has not yet been sequenced. However, OHV-2 was demonstrated to contain one open reading frame (ORF) which was homologous to AHV-1.

6.2 SEQUENCE COMPOSITION AND ANALYSIS

The G+C content of each of the three clones was analysed using the GCG programme COMPOSITION. The three clones, ATT-1, VIR-1 and VIR-2 had G+C compositions of 44%, 47% and 45% respectively. These percentages were compared to the G+C contents of other rhadinoviruses as demonstrated in figure 6.1. The limited sequence data generated suggests that AHV-1 unique DNA is slightly deficient in G+C. The dinucleotide frequencies showed a markedly lower CpG frequency than would be expected from the mononucleotide compositions as illustrated in figure 6.2a. (The predicted CpG frequency was calculated by multiplying the probability of a C by the probability of a G). The difference between the actual CpG frequencies and the predicted CpG frequencies is approximately three-fold for all three clones. This CpG depletion has been observed in the gammaherpesviruses EBV, HVS and *murid herpesvirus* 4 (Honess *et al* 1989, Efstathiou *et al* 1990a). Honess *et al* (1989) observed a corresponding increase in TpG and proposed that this could be the result of the mutagenic effects of DNA methylation during latency in a dividing cell. A slight increase in the frequency of TpG is observed in the C500 DNA sequence, as illustrated in figure 6.2b.

Figure 6.1

G+C content of rhadinoviruses

rhadinovirus	G+C content (complete genome)	G+C content (unique L DNA)
* *Ateline herpesvirus 2 (herpesvirus ateles)*	47.1%	37.5%
* *Saimirine herpesvirus 2 (herpesvirus saimiri)*	45.4%	35.8%
** *Murid herpesvirus 4*	-	45%
Alcelaphine herpesvirus 1	***50%	ATT-1 44% VIR-1 47% VIR-2 45%

References
* Fleckenstein and Mulder 1980
** Efstathiou *et al* 1990b
*** Bridgen *et al* 1989

Figure 6.2

Deviations from expected dinucleotide frequencies

Figure 6.2a
Comparison of predicted and actual CpG frequency of the C500 clones

Clone	Actual CpG frequency	Predicted CpG frequency
ATT-1	63	185
VIR-1	61	186
VIR-2	61	170

Figure 6.2b
Comparison of predicted and actual TpG frequency of the C500 clones

Clone	Actual TpG frequency	Predicted TpG frequency
ATT-1	279	225
VIR-1	278	202
VIR-2	255	200

6.3 DNA DATABASE SEARCHING

Extensive database searching was carried out using the software and databases available on the SEQNET molecular biology facility. ATT-1, VIR-1 and VIR-2 were compared to sequences in the EMBL database using the GCG FASTA DNA and protein database searching programme (Lipman and Pearson 1985). No significant homology was observed between the cloned DNA and sequences stored in the database.

6.4 IDENTIFICATION OF PROBABLE CODING REGIONS IN ATT-1, VIR-1 AND VIR-2

Potential coding regions of the three clones were predicted using TESTCODE. This programme identifies protein coding sequences by plotting a measure of the composition at every third base, based on Fickett's testcode statistic (Fickett 1982). Significant predictions are made for window sizes of 200bp or more, therefore in all cases the window size selected was 200bp. TESTCODE outputs were prepared for all three clones, as exemplified in figure 6.3. TESTCODE identifies codons for methionines, stop codons and predicts with a 95% confidence limit that any sequence above the line at 9.5 is coding DNA, and conversely a confidence limit of 95% that any sequence below 7.5 is non-coding. A prediction is not made for the region of DNA between these two parameters. Hypothetical peptides were identified from the regions of ATT-1, VIR-1 and VIR-2 above the 7.5 TESTCODE non-coding 95% confidence limit. The minimum size of the hypothetical peptides was 80 amino acids, and the sequences were not required to begin with a methionine to allow for the presence of introns. Nine hypothetical peptides were coded for by one strand of sequence (see figure 6.4), with only one being encoded by the opposite strand. The sequence in appendix 3 is presented with the strand which encodes the nine hypothetical petides beginning with base 1, although this is the opposite orientation from the previous descriptions of the clones in chapters 3, 4 and 5. The single ORF in the "non-coding" strand did not exhibit homology to any protein sequences in the database and will not be discussed further.

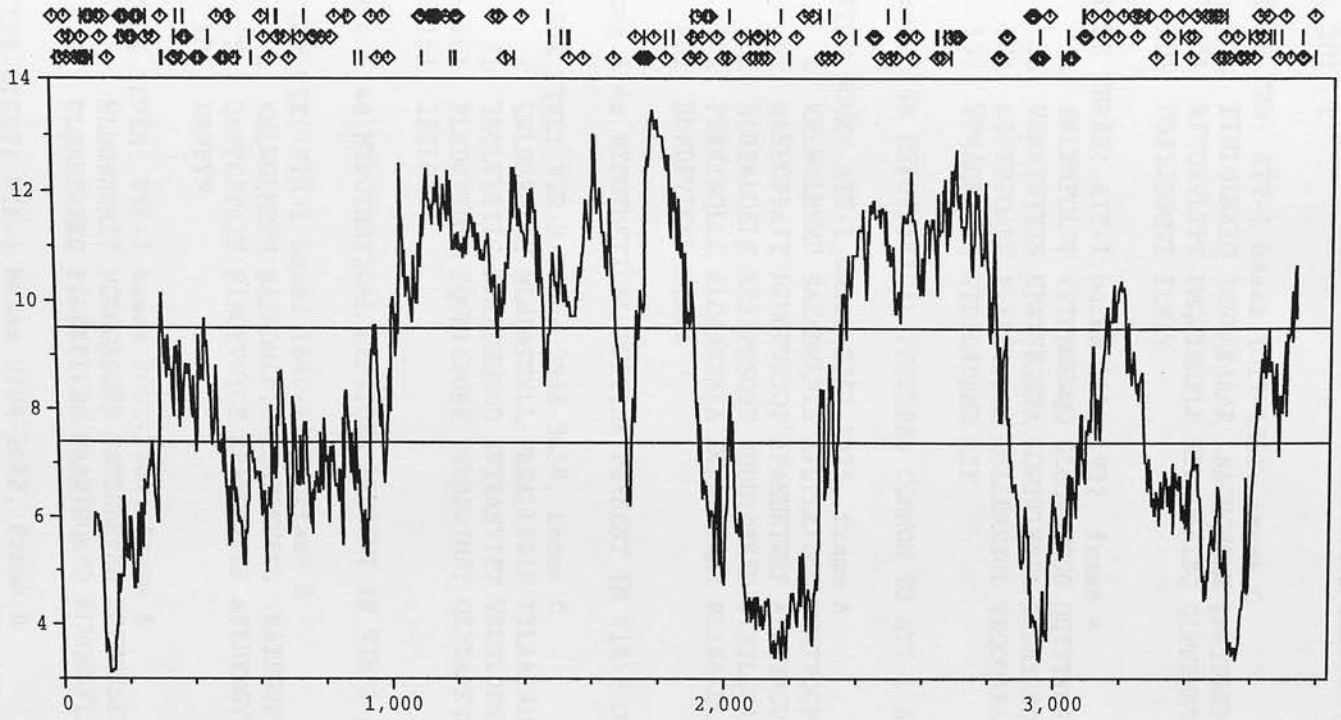


Figure 6.3 TESTCODE of ATT-1

◇ - stop codons | - methionines

Figure 6.4

Figure 6.4a HYPOTHETICAL PEPTIDES COMMON TO ALL THREE C500 CLONES

PEPTIDE ONE: ATT-1 bases 845-1747, frame B

```
  1  SHLDTVSAVC TKERCPNKTV FSQYGARQRY VLPEVLLEAS EPQSTLAYGS
 51  PDISSLLRDS SSTQENTDEE SGPCCSKTLS PGVPQPQSEY DPSPTSPPDS
101  DTESCDSQVR PESTDSDTHA EDDDVPEAPQ AASQTQPTTT QETQSCCSVY
151  NPATTQTGFY YNQQSSDNFA SRIDASATST SYGYPGPVTN HGFSTSVFSH
201  TVPTTATGSS SCIKICMGVV SRRPHTAAMW GVTQMLMASL LVQVPVTTPV
251  SLRPAAAAPL CSGPLCFHRP AHLLPAAKFL IISAPSLIFL AVLQFLRPSS
301  L
```

PEPTIDE TWO: ATT-1 bases 1017-1361, frame C

```
  1  EIHLQLKKIQ TKNLARAAAR LSAQVCHSLN LNMTPARPLH LIRTLSHVIA
 51  RFVQKVLTPT PMQKTMMFLK HPKQQAKLSQ QLPKRHRVAA VSTIQLPLRR
101  DSTITSSQVI ILPLV
```

PEPTIDE THREE: ATT-1 bases 1414-1932, frame A

```
  1  SRLFNFCLFP YSTHNGHWQQ QLHQNMYGGG QQTTTYGSYV GGYSDANGQS
 51  VGASTSYYTS KPATSRSSTA LQWTTLFPPA SSSASSSQVS DYFSPFPYFP
101  GSSTVPQAFE PLAPSTPSLL DELLDRDSGL VSQQQAPAPP QNDQGGPPQY
151  VPVAQEQQQS STDPLSDEMR RIF
```

Figure 6.4b HYPOTHETICAL PEPTIDES COMMON TO ATT-1 AND VIR-2

PEPTIDE FOUR: ATT-1 bases 2281-2925, frame A

```
  1  KHLWGLMHKH SATADMLPIS AFSPLKFLGG VVATPQGPVL WDATTTPVPF
 51  QAFPCPPAIT PINQFVNYQP APAMDKIWRP AFEDYRGVTM RPALPEVPET
101  SLRPALQELP EPSSPQSQSS VDDDTDSKED VTETLECAQA LTDLKWGTVD
151  PSEKTSQLLL ALQAARPAFA EGKGTAIRAY RLIEAKDLGK LVVRGPKKDK
201  RDPDFYIKKV PLHFS
```

Figure 6.4c HYPOTHETICAL PEPTIDE PRESENT IN VIR-2 ONLY

PEPTIDE FIVE: VIR-2 bases 2988-3458, frame C

```
  1  LWLGSEMLAE MLWPAVNMLL PRKALLVDIF FILAATNLMI AAFALGCLAF
 51  YKQLVYITIG NLTFPHQSGD EVIRAMYIPP VNDSVDFNPG FRLSWLNTLS
101  PLSDGPYDSW SQCEICPGRF VSKRPVIMSP QRPIAFKTAS LPVRIFLNVF
151  IFILHRI
```

Figure 6.4d HYPOTHETICAL PEPTIDES PRESENT IN VIR-1 ONLY

PEPTIDE SIX: VIR-1 bases 1646-1963, frame B

```
  1  KHLWGLMHKH SATADMLPIS AFSPLKFLGG VVATPQGPVL WDATTTPVPF
 51  QAFPCPPAIT PINQFVNYQP APAMDKIWRP AFEDYRGVTM RPAKTHYSCS
101  PRVPYN
```

PEPTIDE SEVEN: VIR-1 bases 2001-2264, frame A

```
  1  NLDGPDSMST PGGGGASGRS PNGLGARPKD KGPKGKSPPK CGPAGKTPPK
 51  DLGARPKEKG PKGKSPPKDG PAARTHQKTQ ELVQKKFL
```

PEPTIDE EIGHT: VIR-1 bases 2204-2482, frame B

```
  1  RRACSKNPPK DPGARPEKVP LAFPGPGPVD TNPSRHDVIE AAPGDENPYK
 51  KMWLPPGTKP PCPTPLWTPV RSCYLVVQTP HGTEGKEYIE ILK
```

PEPTIDE NINE: VIR-1 bases 2341-2779, frame A

```
  1  ESVQKNVAAS RNEATMPYTL MDPCAVLLSS GSNAPWDRRK GIHRNTKVKT
 51  IGTPVTPVYL TISDDDLDDA KVLQEVVLKS LLFQAEILQV EAGKNLLLYH
101  LGTTNPASPE LCA
```

6.5 PROGRAMMES USED FOR PROTEIN DATABASE SEARCHING AND PEPTIDE ANALYSIS

6.5 1 Protein database searching

Homologous protein sequences are more likely to be conserved than the genes which encode them since selection acting at the protein level may conserve the sequence, whilst silent mutations, or mutations resulting in conservative substitutions may result in more substantial changes in the DNA sequence. Differences in the G+C composition and codon usage between different species may cause substantial divergence at the DNA level while similarity at the protein level is maintained.

Protein database searching using the nine hypothetical peptides was achieved using the PROSRCH protein sequence database searching programme (Coulson *et al* 1987), which uses the "best local similarity" algorithm of Smith and Waterman (1981). When these searches were undertaken 42,215 polypeptides comprising some 12,368,860 amino acids were present in the database. The search programme collects alignments between the input sequence and the database sequences and produces a score based on the system of Dayhoff *et al* (1978), which looks at accepted point mutations (PAMs). PAM is a measure of divergent evolution whereby one PAM = 1 substitution per 100 aligned residues. The PAM value used when searching the database for similarities to the C500 hypothetical peptides was 250, to allow for distant matches.

The SWEEP matrix optimal alignment programme was also used to compare the hypothetical peptides to sequences in the database. The SWEEP program uses the algorithm of Lipman and Pearson (1985) and scans each probe sequence twice. Optimal alignment is achieved following a modification of the Needleman and Wunsch (1970) algorithm. The default matrix is 250 PAM.

6.5.2 Statistical analysis of the degree of homology between two peptides

PROSRCH and SWEEP were used to compare the hypothetical peptides to other database entries. Statistical analysis of homology was assessed using GCG BESTFIT, which uses the "best local similarity" algorithm of Smith and Waterman (1981), combined with the score system of Dayhoff *et al* (1978), in a similar manner to PROSRCH. Statistical analysis was achieved by rearranging one of the sequences randomly one hundred times, whilst maintaining length and composition, then realigning it to the first sequence. The results were expressed as quality (of the

match) +/- the standard deviation calculated by BESTFIT. Any homology above the 95% confidence limit (i.e. the average match plus two standard deviations) was regarded as being significant.

6.5.3 Comparison of peptides according to their secondary structure

GCG PEPTIDESTRUCTURE was used to predict the secondary structure of the hypothetical peptides by assessing hydrophobicity according to the method of Kyte and Doolittle (1981) and secondary structure according to Chou and Fasman (1978), with the results being displayed graphically using PLOTSTRUCTURE. The output form of PLOTSTRUCTURE selected was a two-dimensional, or "squiggly", plot. The predicted structural features of peptides were demonstrated by sine waves for α-helices, sharp saw tooth waves for β-sheets, 180° turns for β-turns and dull, saw tooth waves for coils. The PLOTSTRUCTURE outputs described in this chapter were overlaid with symbols demonstrating local hydrophobic and hydrophilic regions within the peptides. In addition, peptide 5 was also overlaid with symbols demonstrating antigenicity. The antigen index was calculated by summing several weighted measures of secondary structure as described by Jameson and Wolf (1988).

The GCG programme MOTIFS was also employed when looking at secondary structure; this programme looks for consensus sequences which are consistent features of functional domains.

6.6 RESULTS OF ANALYSIS OF THE C500 HYPOTHETICAL PEPTIDES

6.6.1 Introduction

The nine hypothetical peptides were each compared to sequenced peptides using PROSRCH and SWEEP. Two-dimensional plots were also prepared for each of these peptides, and are included in the remainder of this section where appropriate. Peptides 2, 3, 6, and 7 all demonstrated no statistically or structurally significant homologies to the proteins in the database. The remaining 5 peptides demonstrated degrees of homology to previously sequenced proteins as discussed in the remainder of this section, with emphasis on homology to other γ-herpesviruses.

6.6.2 Analysis of peptide 1

The SWEEP output for peptide 1 demonstrated limited homology between this peptide and a protein sequence from the γ-herpesvirus *Bovine herpesvirus*-4

(BHV-4). The BHV-4 protein exhibited 39% similarity and 18% identity to peptide 1 over 319 residues (i.e. encompassing the entire length of peptide 1). The quality of the match was 108.7, average 100.8 +/-3.0.

The BHV-4 protein exhibited homology to EBV BRLF1 ORF (van Santen 1993). This EBV ORF encodes an immediate early protein involved in transcription regulation as an activator. The BRLF1 ORF is encoded between nucleotides 103,369 and 105,183 in the EBV genome. The HVS homologue of EBV BRLF1, HVS ORF 50, is located between bases 70,798 and 72,402 in the HVS genome (Albrecht *et al* 1992). The HVS ORF 50 and EBV BRLF1 proteins exhibit highly significant homology to each other, as the quality of the match was 230, average 164 +/- 5.3.

The SWEEP output for peptide 1 did not list matches to either EBV BRLF1 or HVS ORF 50, however, peptide 1 was compared with each of these peptides using BESTFIT. Peptide 1 demonstrated 40% similarity and 24% identity to BRLF1, with the quality being 107.2, average 102.0 +/- 3.3, while 35% similarity and 18% identity was observed between peptide 1 and ORF 50, with the quality being 96.1, average 97.7 +/- 3.8. The regions of EBV BRLF1 and HVS ORF 50 which were matched to peptide 1 by BESTFIT were compared to each other. This match exhibited 42% similarity but only 9% identity and was not statistically significant as the quality was 99.8, average 95.1 +/- 3.0. This region of these two peptides, in common with peptide 1, demonstrated no consensus sequences for structural motifs as listed in the MOTIFS package.

The ORFs which encoded the BHV-4 peptide, EBV BRLF1 and HVS ORF 50 consisted of 552, 605 and 535 amino acids respectively, compared to the 301 residues of peptide 1. These larger peptides were all encoded from a single ORF, suggesting that if the homology to peptide 1 was significant, peptide 1 should be part of a larger ORF. The location of EBV BRLF1 and HVS ORF 50 is more central in their respective genomes than is the terminal location of peptide 1. Co-linearity has recently been established between HVS and AHV-1 by sequencing small regions interspersed throughout the AHV-1 genome (A. Ensser, personal communication). The AHV-1 equivalent of HVS ORF 50 has not yet been sequenced. Regions 5' and 3' to the aligned position of HVS ORF 50 in the AHV-1 genome have been sequenced. These sequences suggest that an AHV-1 equivalent to HVS ORF 50 is more likely to be situated in a similar location to that of HVS ORF 50 than the terminal location of peptide 1.

In conclusion, the statistical significance of the match between peptide 1 and the BHV-4 peptide was only just above the 95% confidence limit (i.e 108.7 compared to 106.8), and when searching 42,215 peptides a random match of this quality is possible.

6.6.3 Analysis of peptide 4

The PROSRCH output from peptide 4 demonstrated a degree of homology to an EBV tegument protein, and to the EBV EBNA1 protein, which is involved in the EBV latent cycle (Baer *et al* 1984, Klein 1989), but neither match was significant. No structural features were common to peptide 4 and either EBV protein, therefore the matches were regarded as random occurrences.

6.6.4 Analysis of peptide 5

The PROSRCH output from peptide 5 demonstrated homology to an EBV probable membrane antigen of 140kD which functions as a tegument protein. BESTFIT analysis demonstrated 45% similarity and 19% identity between these peptides with the quality being 56.1, average 59.1 +/- 1.9. This match was not statistically significant, nor was the match to the HVS 140kD membrane antigen equivalent. The structures of the PROSRCH "matched" regions of peptide 5 and the EBV protein were compared (figures 6.5a and 6.5b). Both peptides demonstrated a hydrophobic region of 15-20 amino acids at the carboxy terminus of the "homologous region", preceded by approximately 40 residues structured as β-turns. The overall structural similarity of the two peptides was lost upstream of this point. The functional significance of the region was assessed using MOTIFS, but no known functional motifs were present in this region of either peptide.

The 140kD EBV tegument protein comprises 1318 amino acids, compared to the 157 amino acids of peptide 5. Peptide 5 may be an exon of a larger protein, as it is encoded close to the terminus of the VIR-2 clone, hence further exons may be encoded downstream.

The limited structural similarity of these two peptides may suggest a common ancestor for that particular region. Alternatively the match may have occurred randomly.

Figure 6.5a PLOTSTRUCTURE of peptide 5

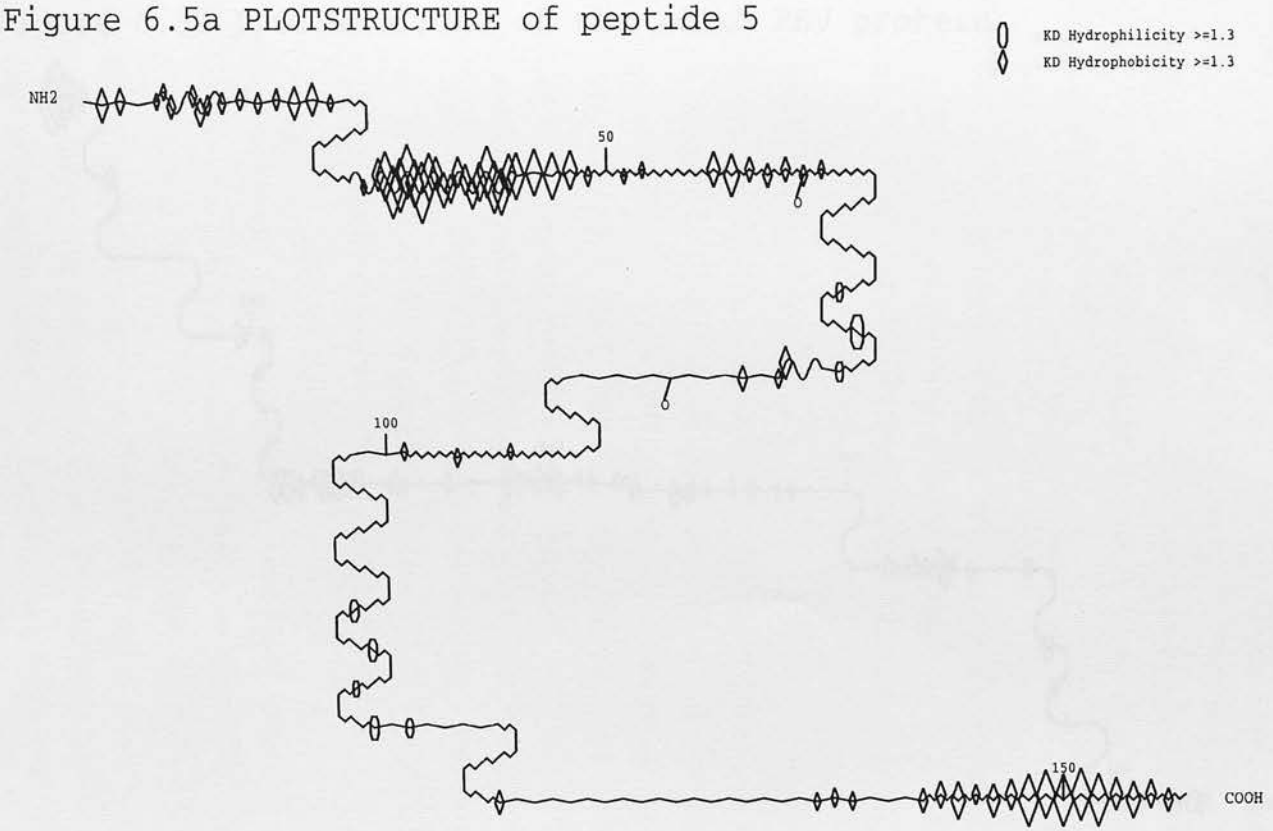
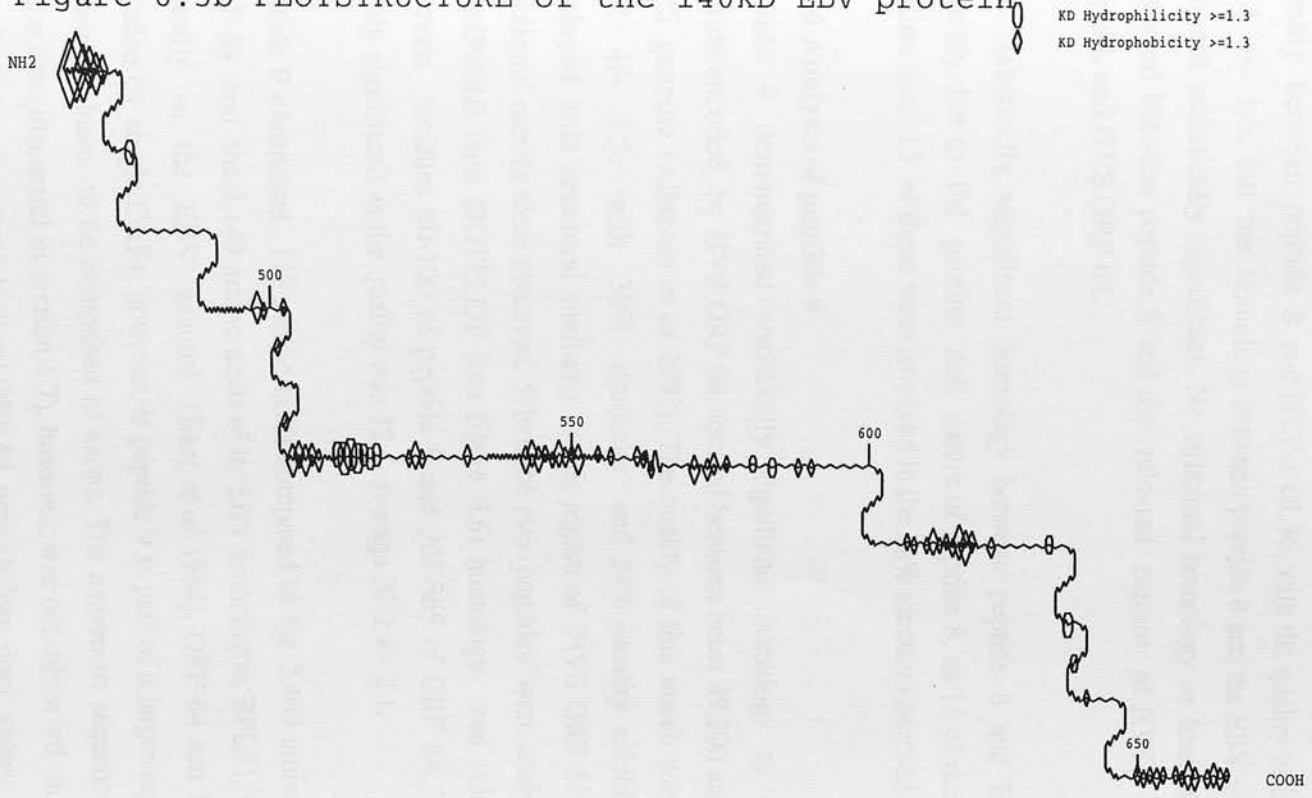


Figure 6.5b PLOTSTRUCTURE of the 140kD EBV protein



6.6.5 Analysis of peptide 8

Peptide 8 demonstrated homology to a herpes simplex virus 1 (HSV-1) protein UL36 (McGeoch *et al* 1988) by PROSRCH. This HSV-1 protein was related to an EBV tegument protein, HVS ORF 64, an equine herpesvirus 1 protein and a varicella zoster virus protein. BESTFIT demonstrated 42% homology and 32% identity between peptide 8 and HSV-1 UL36, with the quality being 43.9, average 39.1 +/- 1.6, but the homology between peptide 8 and the EBV and HVS proteins was not statistically significant. No structural homology or functional motifs were observed between peptide 8 and the relevant regions of HVS-1 UL36, the EBV protein, and HVS ORF 64.

The statistically significant homology between peptide 8 and HSV UL36 was possibly due to the proline rich nature of peptide 8, as 18 of the 93 residues are proline and 13 of these were involved in the 32% identity observed.

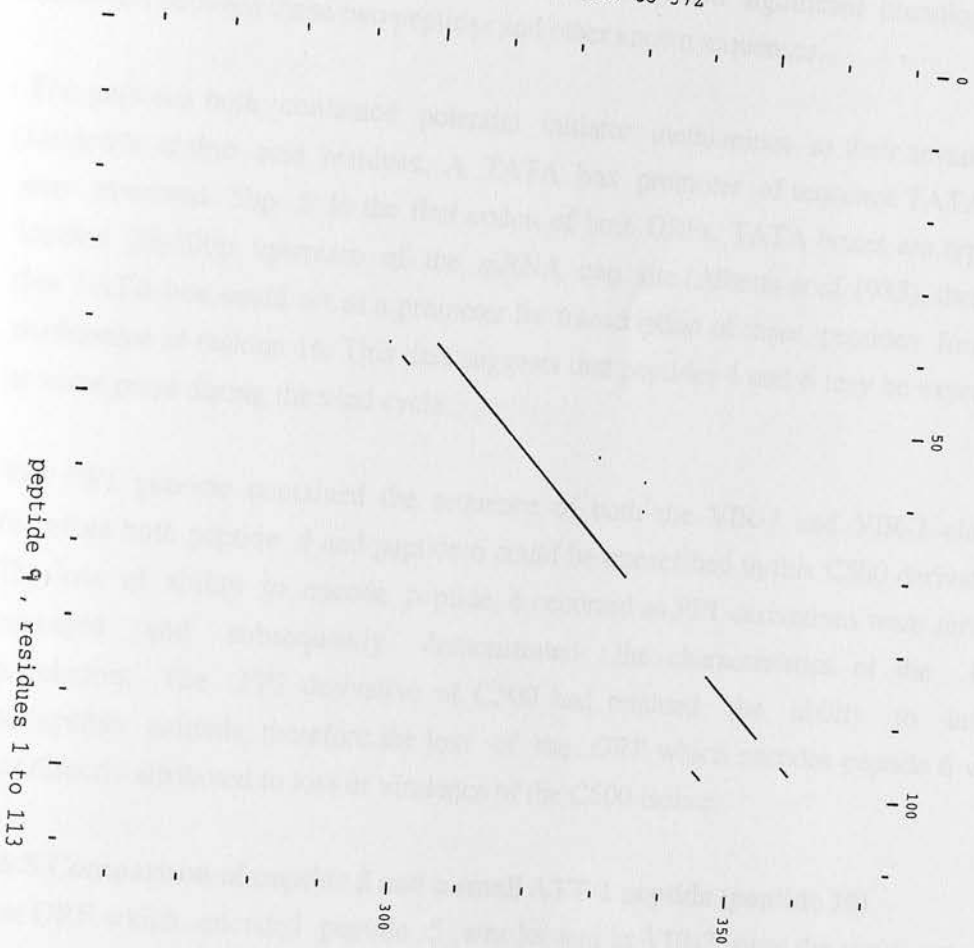
6.6.6 Analysis of peptide 9

Peptide 9 demonstrated statistically significant homology to the hypothetical peptide encoded by HVS ORF 64, located between bases 89,200 and 96,606 in the HVS genome (Albrecht *et al* 1992). The quality of this match was 49.8, average 45.2 +/- 1.5, with 38% similarity and 24% identity exhibited. Peptide 9 displayed little structural similarity to this region of HVS ORF 64, and no known functional motifs were observed. When the two peptides were analysed using GCG COMPARE then DOTPLOT (see figure 6.6) homology was clearly observed between residues 40-100 of peptide 9 and 305-365 of ORF 64. This match was highly significant as the quality was 32.4, average 20.2 +/- 2.1.

Peptide 9 comprises 113 amino acids compared to the 2,469 amino acids of HVS ORF 64 and the 3,149 amino acids of its EBV homologue, BPLF1, which is located centrally in the EBV genome (Baer *et al* 1984). ORF 64 and BPLF1 are both encoded by single ORFs, however, if peptide 9 is part of a larger equivalent protein it would have to be comprised of exons. The consensus sequence for a 3' exon splice site (discussed in section 6.7), however, was not observed in peptide 9. The homology between peptide 9 and ORF 64 nevertheless does appear to be genuine, and perhaps the peptides share a common ancestor and function.

Peptide 9 was encoded by the region of VIR-1 present in all C500 derivatives. This ORF was not altered during *in vitro* propagation of the C500 isolate, but the

Figure 6.6 DOTPLOT of peptide 9 compared with HVS ORF 64



sequence upstream was rearranged in the PP2 and CFA genomes, compared to PP1 DNA, possibly affecting expression. Expression of peptide 9 alone could not be entirely responsible for virulence as the rearrangements which surround this peptide occur in the PP2 derivative, a C500 derivative known to be virulent.

6.6.7 Comparison of peptide 4 and peptide 6

The ORFs which encode hypothetical peptides 4 and 6 share common sequence in their first 93 amino acids, as both ORFs are located over the region of homology loss between VIR-1 and the other two clones. No significant homology was observed between these two peptides and other known sequences.

The peptides both contained potential initiator methionines as their seventh and sixteenth amino acid residues. A TATA box promoter of sequence TATAAAA was observed 5bp 5' to the first codon of both ORFs. TATA boxes are typically located 25-30bp upstream of the mRNA cap site (Alberts *et al* 1983), therefore this TATA box could act as a promoter for transcription of these peptides from the methionine at residue 16. This data suggests that peptides 4 and 6 may be expressed at some point during the viral cycle.

The PP1 genome contained the sequence of both the VIR-1 and VIR-2 clones, therefore both peptide 4 and peptide 6 could be transcribed in this C500 derivative. The loss of ability to encode peptide 6 occurred as PP1 derivatives were serially passaged and subsequently demonstrated the characteristics of the PP2 population. The PP2 derivative of C500 had retained the ability to infect susceptible animals, therefore the loss of the ORF which encodes peptide 6 was not directly attributed to loss of virulence of the C500 isolate.

6.6.8 Comparison of peptide 5 and a small ATT-1 peptide (peptide 10)

The ORF which encoded peptide 5 was located in VIR-2, over the region where homology is lost between VIR-2 and ATT-1. The equivalent ORF in ATT-1 was located between nucleotides 3632 and 3832, and encoded 70 amino acid residues. No stop signal was observed before the end of the sequenced DNA. The sequence of peptide 10 is:

1 LWLGSEMLAE MLWPAVNMLL PRKALLVDIF FILAATNLMI

40 AAFALGLLAF YNAEISDLSA DATPLSPGPP

The first 51 residues of this peptide were common to peptide 5. No significant homology was observed when peptide 10 was compared to other protein sequences in the database.

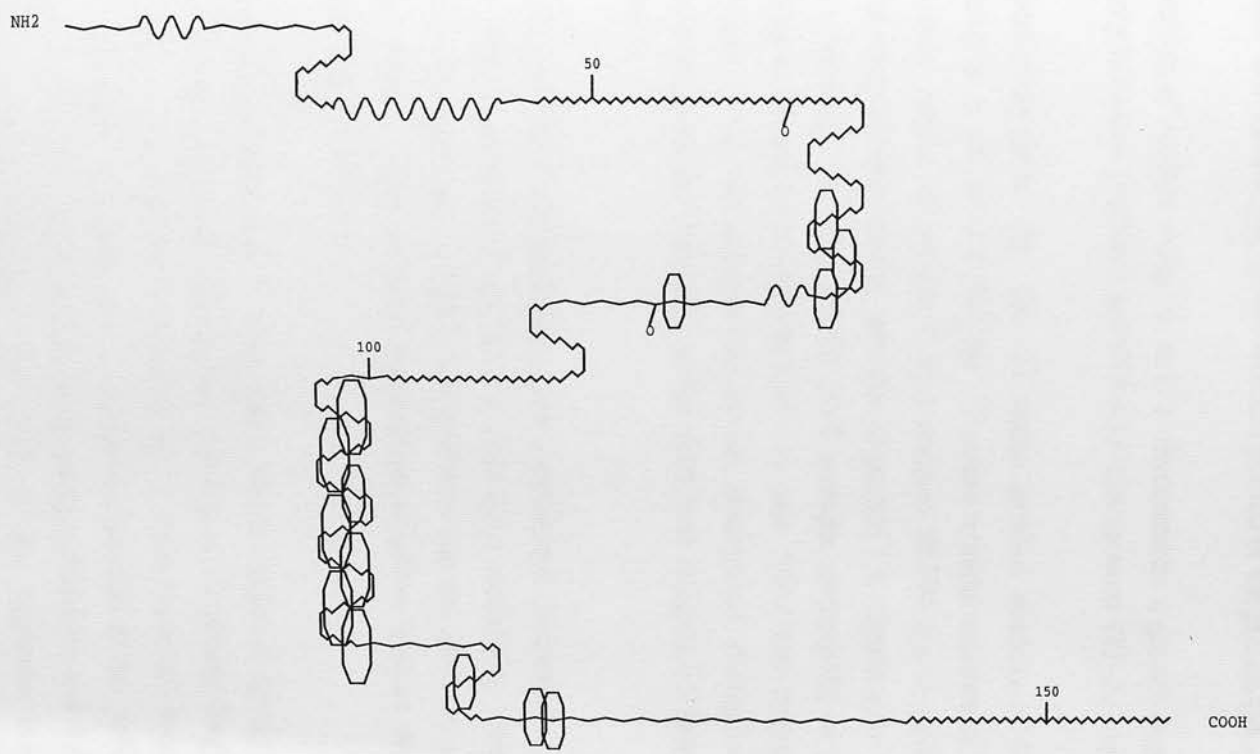
The seventh, eleventh and eighteenth residues of peptides 5 and 10 are methionines. A sequence of TATTTA, an almost exact copy of the SV40 early promoter TATTTAT (Laimins *et al* 1982), was observed 7bp 5' to both ORFs, suggesting that either of the methionines at positions 11 and 18 may be initiator codons. Figure 6.5a shows a 15-20 amino acid hydrophobic sequence at the carboxy-terminus of peptide 5, representing a possible membrane spanning region. The antigenicity plot of peptide 5 (figure 6.7), predicts residues 100-130 to be strongly antigenic. The DNA sequences which encode the hydrophobic and antigenic domains of peptide 5 are absent from the CFA genome. These sequences are lost as C500 virions adapt from resembling the PP2 population to resembling the CFA population, suggesting that peptide 5 may be important in the virulence mechanism of the C500 isolate.

6.7 FURTHER ANALYSIS OF PEPTIDE 1

Peptide 1 was encoded by the largest C500 ORF sequenced, and was common to all three clones. The structure of peptide 1 exhibited a marked polarity in that the amino terminus of the peptide was found to be highly hydrophilic, while the carboxy terminus was highly hydrophobic. Peptide 1 was located close to the terminal repeats of C500, therefore was compared to similarly located peptides in the rhadinovirus prototype, HVS (Albrecht *et al* 1992).

The ORF closest to the 5' terminus of the HVS subgroup A, strain 11 encodes a HVS transformation-associated protein, STP-A11 (Murthy *et al* 1989). An equivalent protein was found in HVS subgroup C, strain 488, STP-C488 (Biesinger *et al* 1990, Jung *et al* 1991). Both STP-A11 and STP-C488 have tumour-inducing activities independent of the rest of the genome. The DNA and amino acid sequences of these HVS peptides demonstrate no identity to other known transforming genes and it has been suggested that they may represent a new class of proteins involved in lymphocyte transformation (Jung *et al* 1991, Beisinger *et al* 1990).

Figure 6.7 Antigenicity PLOTSTRUCTURE of peptide 50



STP-A11 and STP-C488 exhibit limited amino acid sequence homology to each other, but are structurally similar, as both have acidic amino termini and hydrophobic carboxy termini. STP-C488 codes for 18 directly repeated collagen-like motifs of Gly-X-Y (where either X or Y is proline) located centrally in the peptide (Jung *et al* 1991). STP-A11 codes for only 9 of these motifs, and these are dispersed throughout the amino terminal portion of the protein (Murthy *et al* 1989).

Comparison of figures 6.8a, b and c demonstrates a greater degree of structural similarity between peptide 1 and STP-A11 than between STP-A11 and STP-C488.

The isoelectric point for the 32 amino terminal residues of STP-A11 is 3.5, compared to a pI of 4.4 for the 17 amino terminal residues of STP-C488. The comparable region of peptide 1, from residues 50-180, has an isoelectric point of 3.5. The hydrophobic carboxy terminus of peptide 1 is significantly longer than the carboxy terminus of either of the HVS proteins, comprising of 100 residues compared to 26 and 28 residues in STP-A11 and STP-C488 respectively. Peptide 1 contains 6 of the collagen-like repeats, interspersed throughout the peptide. The three peptides are compared in bar chart form in figure 6.9 (adapted from Jung *et al* 1991).

STP-A11 and STP-C488 both begin with a methionine. In contrast, peptide 1 neither begins with a methionine, nor does it code for a methionine close to the amino terminus. In addition, a TATA box promoter was not observed upstream of this reading frame. Lack of these transcription initiation features suggest that this sequence may be an exon.

A consensus sequence for 3' exon splice sites is conserved throughout eukaryotes. The sequence consists of a pyrimidine rich region of variable length (but always more than 10 nucleotides) followed by a short consensus sequence extending only 3 residues upstream and 1 residue downstream of the splice point. This sequence is $(Py)_nNPyAGG$, with Py representing pyrimidines and the final G being the first nucleotide of the exon. The AGG of this sequence is almost invariant (Watson *et al* 1987). The closest match to this consensus sequence was located between nucleotides 862-882 of ATT-1 (TCTCTGCTGTTTGCACCAAGG) which encode residues 6-13 of peptide 1. The first coding base of the exon would be 882, corresponding to amino acid 13 (see figure 6.4). The region encodes the criteria for a 3' splice site with the exception of base 879, which should be a

Figure 6.8a PLOTSTRUCTURE of peptide 1

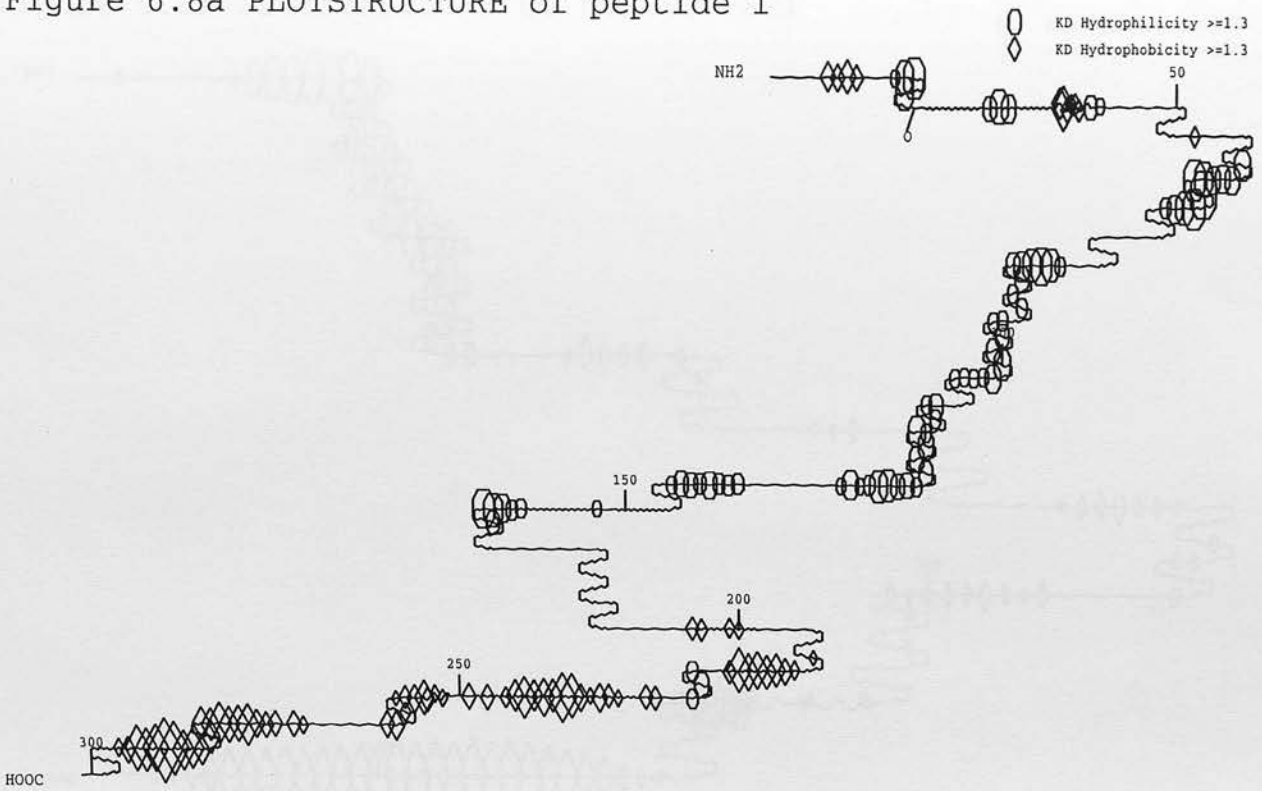


Figure 6.8b PLOTSTRUCTURE of HVS STP-A11

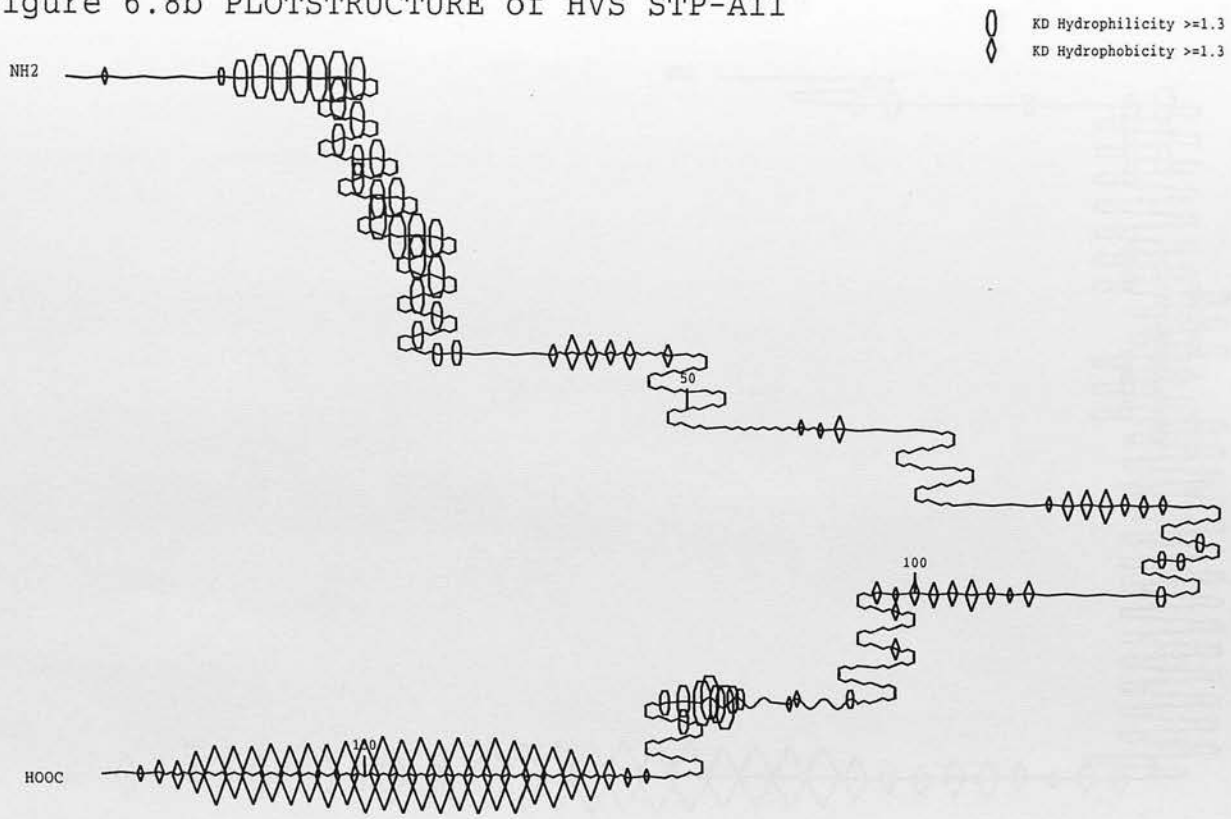


Figure 6.8c PLOTSTRUCTURE of HVS STP-C488

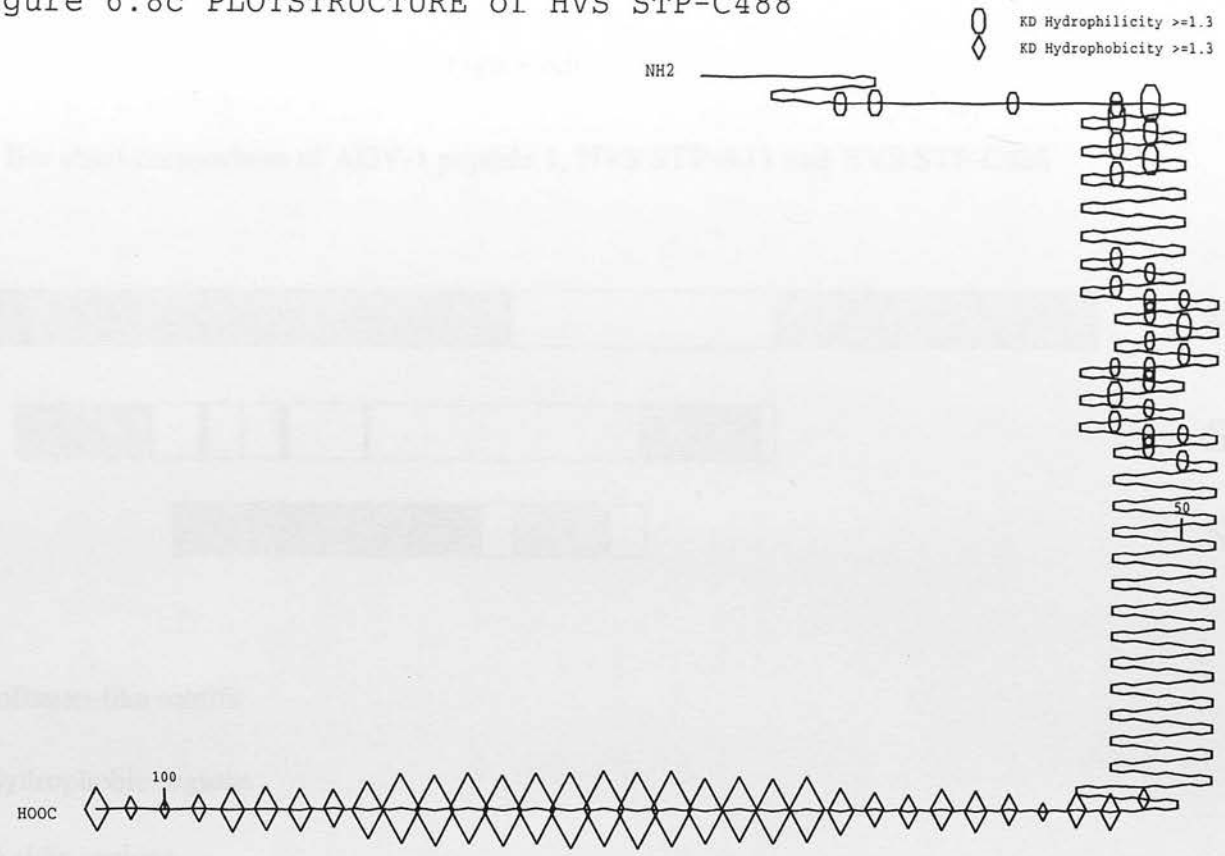
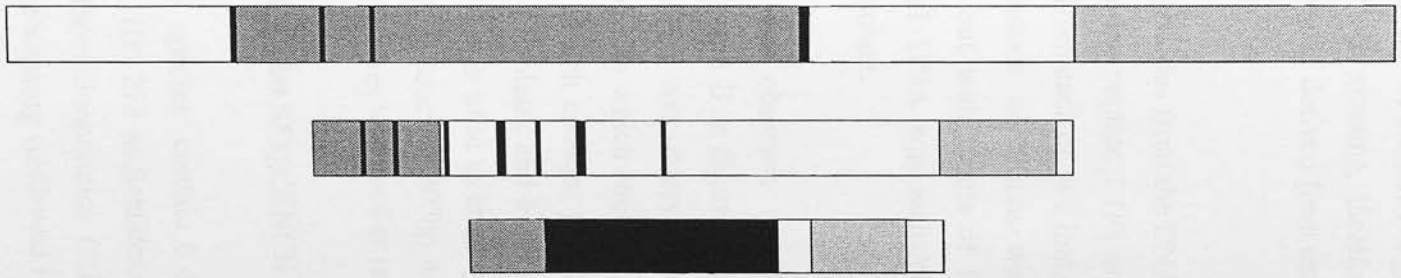


Figure 6.9

Bar chart comparison of AHV-1 peptide 1, HVS STP-A11 and HVS STP-C488



Key

Collagen-like motifs

Hydrophobic regions

Acidic regions

(adapted from Jung *et al* 1991)

pyrimidine but is an adenine (reviewed in Watson *et al* 1987, Jackson 1991). In conclusion, amino acid 13 of peptide 1, glutamine, may be the first amino acid of this exon.

6.8 AN OHV-2 HOMOLOGUE OF PEPTIDE 1

The ORF which encoded peptide 1 was observed in the PP1, PP2 and CFA derivatives of C500. OHV-2 is closely related to AHV-1, therefore a peptide 1 homologue was looked for in the OHV-2 genome. The OHV-2 DNA has not yet been isolated as an intact genome, therefore OHV-2 DNA was isolated from immortalised T-lymphocytes derived from cattle affected with OHV-2 induced MCF (Reid *et al* 1989).

A PCR using oligonucleotides from the DNA sequence located to the 5' and 3' ends of the ORF which encodes peptide 1 (P8 and P11 in appendix 3) was carried out using 500ng of DNA from each of two immortalised cell-lines (kindly prepared by Miss I Pow). The reaction conditions were as described in 2.8.20 F. Identical reactions were carried out with inputs of 100ng of PP1, CFA and WC11 DNA respectively. The WC11 DNA was included to ascertain whether peptide 1 was present in this AHV-1 isolate.

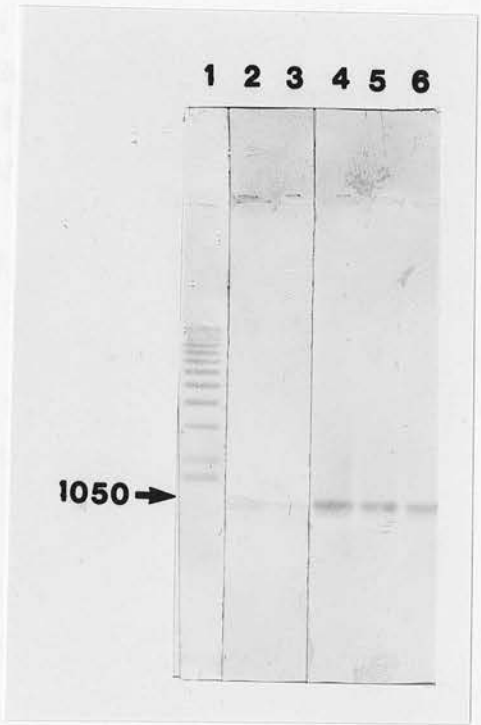
A product of 1050bp was observed in all five PCR reactions. These products were probed with fragment B in figure 3.12, which is located internal to the PCR primers used, and homology was observed (see figure 6.10), demonstrating that the region of the C500 genome which encodes peptide 1 is also present in WC11 and in OHV-2. The region which encodes peptide 1 appears to be highly conserved between the two AHV-1 isolates and between AHV-1 and OHV-2. An additional PCR using P3 and P14 was used to amplify a 400bp fragment internal to P8 and P11 (see appendix 3). A product of 400bp was observed when the DNA from the OHV-2 immortalised cell-lines was used as target (data not shown).

6.9 SEARCHING THE C500 SEQUENCE FOR THE PRESENCE OF U RNAS

Cells of all vertebrate species contain 6 different small nuclear RNAs (U1-U6), ranging in size from 100-215 nucleotides, which are complexed with proteins to form small nuclear ribonucleoproteins (snRNPs). snRNPs are involved in pre-mRNA splicing and processing (reviewed in Maniatis and Reed 1987). Marmoset

Figure 6.10

Southern blot of AHV-1 and OHV-2 PCRs



Track 1 1kb ladder
Track 2 OHV-2 cell-line 1 PCR
Track 3 OHV-2 cell-line 2 PCR
Track 4 PP C500 PCR
Track 5 CFA C500 PCR
Track 6 WC11 PCR

Tracks 2-6 were probed with fragment B from Figure 3.12

T lymphocytes transformed with HVS contain the first virally encoded U RNAs (known as HSURs) to be identified. Their sequences are related to each other but are distinct from any previously characterised cellular U RNA (Lee *et al* 1988), and are encoded adjacent to STP-A11 in HVS strain 11 (Albrecht *et al* 1992). These HSURs contained three consensus sequences (Albrecht and Fleckenstein 1992):

sm-site AUUUUUG
stem GGCNCTGG
stem CCAGNGCC

C500 homologues of these sequences were searched for using BESTFIT to analyse each consensus sequence individually. Near-perfect matches were observed to each of these sequences, but the matches were too distant from each other to be coding for HSUR equivalents, since HSURs range in size from 75 to 140 nucleotides (Murthy *et al* 1986).

6.10 DISCUSSION

The sequence of the C500 genome represented by the three clones, ATT-1, VIR-1 and VIR-2, was compared to all other DNA sequences entered in the EMBL DNA database. No significant homology was observed.

The G+C content of the cloned DNA was consistent with other rhadinoviruses, and the CpG ratio was consistent with γ-herpesviruses as a whole. The short repeated regions observed were also a common feature of γ-herpesviruses.

Exhaustive analysis of hypothetical peptides both via amino acid sequence and structural similarity revealed very little homology to sequences in the database. The most statistically significant homology observed was between a short region of peptide 9 and HVS ORF 64, however the biological significance of this homology is not known. The best structural homology observed was between peptide 1 and the HVS transforming protein STP-A11. It is not known at present whether peptide 1 is expressed in active AHV-1 or OHV-2 infection, nor if this hypothetical peptide has transforming ability. The pathogenesis of MCF is not yet fully understood, however, AHV-1 and OHV-2 both have the ability to immortalise T-lymphocytes (Wilkinson *et al* 1992, Reid *et al* 1989a). Peptide 1 may function in a similar manner to STP-A11 and STP-C488 and may have the potential to transform lymphocytes. However, the presence of two copies of peptide 1 in the attenuated C500 genome (see chapter 4) suggests that peptide 1 may not be the AHV-1

equivalent of the HVS STPs, or that additional factors are required in order for peptide 1 to be functional.

The sequence encoding peptide 6 was known to be genomically altered during *in vitro* propagation of C500 virus, such that the ORF encoding peptide 6 was present in the PP1 derivative but absent from the PP2 and CFA derivatives. However, the first 76 residues from the initiator methionine were included in peptide 4, which was present in all three C500 derivatives. The loss of peptide 6 did not appear to be directly related to loss of virulence of the C500 molecule.

The sequence analysis described in this chapter identified a potential coding region which may be involved with virulence, i.e. the ORF for peptide 5. The final 106 amino acids of this sequence encodes the potential antigenic and membrane spanning regions of the peptide, and is absent from the CFA genome. This sequence is lost from the C500 genome, implying that peptide 5 may be a factor required for virulence.

In conclusion, the region of the C500 genome sequenced showed little homology to other sequences in the DNA or protein databases. The ORF which encodes peptide 5 presents a potential virulence factor, which may act in conjunction with peptide 1.

Chapter 7
Discussion

7.1 INTRODUCTION

The C500 isolate of AHV-1 is altered on serial passage in cell culture, resulting in a loss of virulence. The aim of this project was to identify and characterise the region of the C500 genome correlated with virulence. This chapter discusses the combined results described in chapters 3, 4, 5 and 6.

7.2 THE BIOLOGICAL SIGNIFICANCE OF THE GENOMIC ALTERATIONS OBSERVED

Following isolation and for the first few passages *in vitro* the C500 isolate of AHV-1 was cell-associated and capable of producing MCF on inoculation into susceptible hosts. On additional passage, the virus became cell-free, retaining virulence, until finally it became attenuated (and cell-free). Molecular comparison of viral DNA using the restriction endonuclease Sma I demonstrated consistent profiles defining stages in this attenuation process. The initial population observed during *in vitro* propagation was PP1. The PP1 derivative represented virulent cell-associated C500 virions. The PP1 derivative characteristics were lost on further passage, and the virions observed exhibited characteristics of the PP2 population. Although the PP2 derivative was described as cell-associated in chapter 5, PP2 DNA was also recovered from cell-free virions in the supernatants of PP2 infected cultures. The PP1 and PP2 genomes were not observed together in any of the C500 viral preparations studied, suggesting that the rearrangement of the PP1 genome which results in the PP2 organisation must be extremely advantageous in the artificial environment of *in vitro* culture, as the PP2 arrangement quickly becomes prevalent. One advantage that PP2 virions may have over PP1 virions is that PP2 virions may exist in a cell-free form. The PP1 infectivity *in vitro* is largely cell-associated and infection appears to pass by cell contact, as illustrated by the syncytia in figure 3.1a. The production of a cell-free virus would allow a more rapid spread of infection in the monolayer, and therefore would present a large replicative advantage.

The genomic reorganisation whereby PP2 virions were replaced by virions exhibiting the characteristics of the CFA genome was a more gradual transition, as mixed populations of PP2 and CFA virions (originally defined as CA virus) were observed over several passages (4-10), before the CFA organisation became dominant. The CFA virions were presumably more replication competent *in vitro*

than their PP1 and PP2 predecessors. The 3.8kbp Sma I fragment observed in C500 attenuated virus hybridised to a fragment of similar size in the standard isolate of AHV-1 (WC11), which had also become attenuated as a result of *in vitro* passage (Plowright *et al* 1965). This suggested that the two attenuated derivatives from different isolates were rearranged in a similar manner as a result of serial passage in tissue culture. This strengthens the suggestion that the genomic alterations identified in this study during the process of attenuation are related to loss of virulence.

The proposed sequential rearrangements of the C500 genome (see figure 5.23) illustrate the *in vitro* alterations of the 3' terminal unique region sequence. A second copy of ATT-1 was observed at the 5' end of the CFA unique region, yet VIR-1 and VIR-2 equivalents were not found as close to the 5' termini of PP1 and PP2. The total DNA lost in the two transitions from PP1 to CFA was >3.8kbp, while the size of the ATT-1 clone was 3.8kbp, therefore the length of the unique region is approximately the same for both the PP1 and CFA genomes. Conservation of genome size has been observed in several γ-herpesviruses. The unique region of BHV-4 has been shown to differ by as little as 1kbp (in a 112kbp sequence) in four different strains, suggesting that size constraints are important (Bublot *et al* 1990). The EBV genome length also appears to be strictly conserved as the B95-8 strain has a 12kbp deletion, compared to other strains, but this is apparently compensated for by an increase in the number of copies of the internal repeat unit (Bornkamm *et al* 1980, Hayward *et al* 1980). It has been suggested that a genome length is maintained which is optimal for encapsidation (Miller 1984). In HVS the replication mechanism is believed to be by concatameric replication, with the enzyme responsible for cleaving and packaging individual molecules cutting at random in one of the repeat units. The second cleavage site is then restricted such that overall genome size is conserved (Stamminger *et al* 1987). One HVS strain, SMHI, lost oncogenicity as a result of serial passage in vero cells, generating SMHI-VERO. Genomic analysis revealed that the SMHI-VERO derivative differed from SMHI by a deletion of approximately 6.6kbp at the 5' terminus of the unique sequence. In addition, a fragment of L-DNA was inserted into the H-DNA of SMHI-VERO. This inserted DNA was an inverted repetition of a 5.6kbp region of the 3' end of L-DNA (Daniel *et al* 1980). The effect of this duplication was that genome length was approximately conserved between the SMHI strain and its attenuated derivative. The second copy of the ATT-1 clone in the C500 genome may be present to conserve genome size to allow efficient replication and packaging.

The second copy of the ATT-1 clone observed in the CFA genome may have arisen from a random duplication to conserve genome length. This seems unlikely because, if the duplication was a random event, the different derivatives would be expected to yield different duplicated fragments. The results described in chapter 4 were produced using different preparations of C500 virus, yet the same sequence was always present twice (i.e. ATT-1). This suggests that duplication of the ATT-1 sequence presents a replicative advantage. The attenuated AHV-1 isolate, WC11, also contains a Sma I fragment of 3.8kbp, which is homologous to ATT-1. It has not been determined whether WC11 contains one or two copies of this sequence.

The mechanism which results in duplication of the ATT-1 clone in the CFA genome is not understood. In addition, it has not been established whether the PP2 genome undergoes two sequential rearrangements, with deletion of the 1.2kbp fragment being the first event, and duplication of ATT-1 being the second. Alternatively, the deletion and duplication events may occur simultaneously.

7.3 ALTERED READING FRAMES AS A RESULT OF ATTENUATION

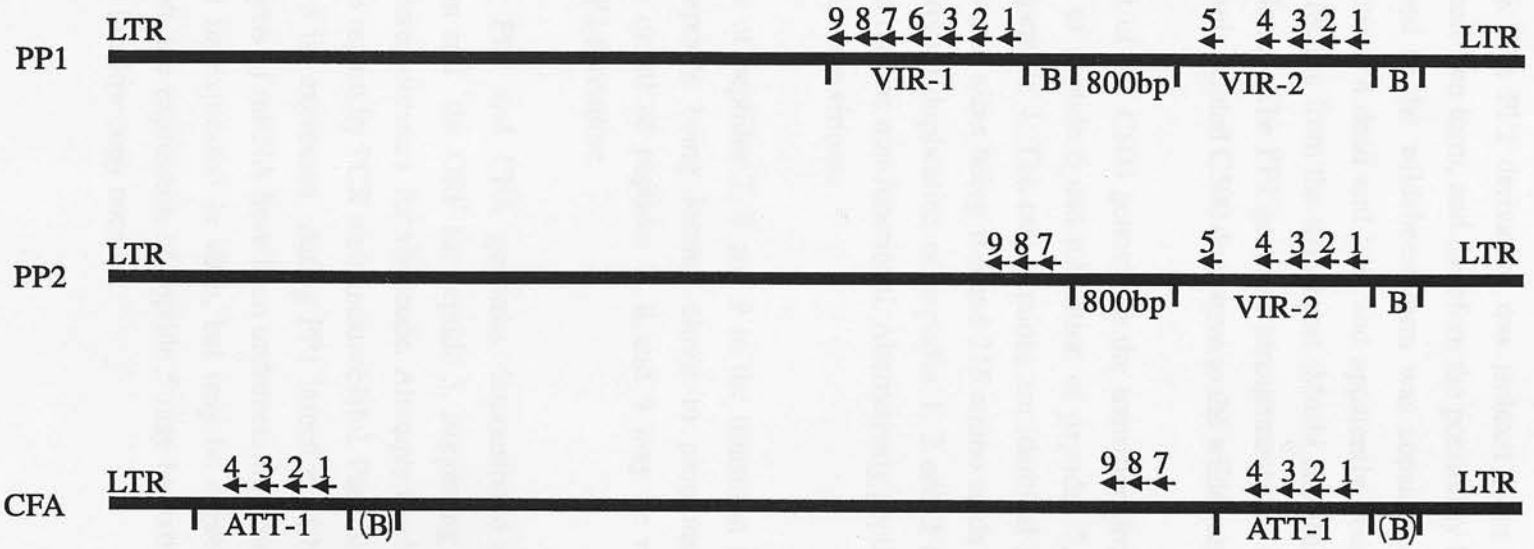
The proposed genomic locations of VIR-1, VIR-2 and ATT-1 in the PP1, PP2 and CFA derivatives were compared to assess which ORFs were altered as a result of serial passage, as demonstrated in figure 7.1. The PP2 genome differs from the PP1 genome by a deletion of fragment B and the region of VIR-1 encoding peptides 1, 2, 3 and 6. The CFA genome differs from the PP2 genome by deletion of the 800bp sequence and peptide 5. In addition, the ability of the C500 virus to cause MCF in rabbits is lost in the transition from PP2 to CFA.

The biological properties of the three C500 derivatives have bearing on the interpretation of the ORFs which are altered as a result of serial passage. The PP1 derivative is cell-associated and represents the first C500 genomic arrangement observed on isolation in cell-culture of virus from experimentally infected rabbits. The PP2 genomic arrangement is found in both cell-associated and cell-free forms, and inoculation of a rabbit with PP2 virus resulted in MCF. CFA virus is largely cell-free, and is avirulent for rabbits.

C500 virus was isolated in tissue culture from a naturally occurring case of MCF in a bovine (Plowright *et al* 1975). It was received in this laboratory following an

Figure 7.1

Diagramatic representation of the C500 altered reading frames



unspecified number of passages as frozen infected rabbit lymph nodes and spleen. This isolate has been maintained as stocks of frozen suspensions of lymph node and spleen cells in 10% DMSO in liquid nitrogen and periodically passaged in rabbits. When isolated from infected rabbits and propagated in tissue culture PP1 viral infectivity was largely cell-associated, although infectious cell-free AHV-1 virus is found in the ocular and nasal secretions of young wildebeest calves (Mushi and Rurangirwa 1981). The PP2 derivative was isolated from infected tissue culture supernatants in a cell-free form, and therefore the possibility that this derivative was more closely related to the wildebeest form was considered. However, the MCF affected animal acts as a dead end host and apparently does not produce cell-free virus on natural infection from the wildebeest (Mushi and Rurangirwa 1981), nor on infection by inoculation. The PP1 genomic arrangement was therefore considered as being the most closely related C500 derivative to the wildebeest virus.

The rearrangement of the C500 genome in the transition from PP1 to PP2 virions results in the loss of peptide 6 and relocation of peptides 7, 8 and 9. Peptide 6 is closely related to peptide 4. The two peptides are identical in their first 93 amino acids, with their overall sizes being 106 and 215 amino acids respectively. Peptide 6 may have arisen from a duplication of peptides 1, 2 and 3 in the evolution of the virus, and therefore may be non-functional. Alternatively peptide 6 may act to inhibit the formation of cell-free virions.

The rearrangement of peptides 7, 8 and 9 in the transition from PP1 to PP2 may result in these peptides being located closer to promoters or enhancers, thus expression of one or all of peptides 7, 8 and 9 may be related to the cell-free properties of the PP2 derivative.

Comparison of the PP2 and CFA genomes demonstrated a deletion of the 800bp unsequenced region and the ORF for peptide 5, suggesting that the 800bp region and/or peptide 5 were necessary for virulence. Attempts to clone this 800bp region, and to amplify the region by PCR were unsuccessful. Preliminary attempts to assess whether peptide 5 is expressed during PP1 infection of bovine thyroid cells by Northern blot analysis of mRNA have been unsuccessful, however it is possible that peptide 5 may not be expressed *in vitro*, but may be expressed in MCF affected animals. In addition, the expression of peptide 5 may be extremely short-lived, and may be transcribed in a low copy number.

The significance of two copies of the ATT-1 clone in the CFA genome with regard to the ORFs duplicated has not been elucidated. Perhaps peptides 1-4 are important for *in vitro* propagation of the virus, and second copies result in up-regulations of 1 or all of peptides 1-4. The observation that peptides 1, 2 and 3 are present twice in the PP1 genome, once in the PP2 genome and twice in the CFA genome is difficult to interpret with regard to the differences between the three C500 derivatives. Perhaps peptide 4, which is present once in PP1 and PP2, but twice in CFA, is involved in increasing the replication competence of CFA virions.

Peptide 1 exhibited structural similarities to the HVS transforming genes STP-A11 and STP-C488 as described in chapter 6. However, the transforming genes were deleted from two HVS strains which exhibited attenuation as a result of *in vitro* propagation. (Daniel *et al* 1980, Koomey *et al* 1984). The presence of two copies of peptide 1 in the CFA derivative, with one of these copies in an identical location to that observed in PP2, was in direct contrast to the absence of the STPs in HVS attenuated derivatives, and suggested that peptide 1 was not functionally similar to the STPs. Peptide 5, although not exhibiting structural homology to the HVS STPs, may be the AHV-1 equivalent. The pathogenesis of MCF is not fully understood (see section 1.7), however, the development of a T cell-line from lymph node cells of a rabbit infected with AHV-1 (Wilkinson *et al* 1992) suggests a transformation-related mechanism, analagous to the development of a lymphoid cell-line from infection with HVS (Schirm *et al* 1984).

Peptides 7, 8 and 9 may be involved in the virulence of the C500 isolate, as the sequence upstream of these peptides is altered between PP2 and CFA genomes, possibly affecting expression.

7.4 POSSIBLE RECOMBINATION FEATURES OF THE C500 SEQUENCE

The genomic rearrangements documented in the previous chapters, and the sequence data generated, were analysed in an attempt to discover recombination mechanisms. Throughout the course of this work various possible recombination mechanisms have been considered. The presence of large, internal, inverted duplications of terminal sequences, correlating with high frequency inversion of the unique sequence elements enclosed in these duplications, has been reported in herpesviruses, and results in either two or four isomeric forms (summarised by Honess 1984). However, the rhadinovirus prototype, HVS, exists in only one

isomeric form (Roizman 1990). The mapping of the AHV-1 isolate WC11 suggested that the genomic arrangement was similar to that of HVS (Bridgen *et al* 1989, Bridgen 1991). The presence of two isomers of C500 would account for the two copies of ATT-1 at either end of the genome, if the entire unique region of the C500 genome was inverted. The observation that the two copies of ATT-1 are present in the same orientation suggests that inversion is not the method of recombination. The mechanism by which the ATT-1 clone is duplicated is not known, but may be linked to its proximity to the terminal repeat units of the genome.

The limited sequence data generated from the three molecular clones exhibited two short regions of repeated DNA, both located within the VIR-1 clone (see section 3.13.3). Second copies of these repeat regions were not observed in the remainder of the region sequenced. The two copies of the 85bp repeat are situated in the region of the breakpoint of the VIR-1 clone which occurs in the transition from PP1 to PP2 genomes, presenting the possibility that this region is somehow involved in the rearrangements observed.

A consensus sequence was observed before the two points of homology loss (see section 3.15). In the transition from PP1 to PP2 (see figure 5.23) the consensus sequence in the VIR-1 clone occurs immediately upstream of the breakpoint in homology with the other two clones (i.e. base 1923). The deleted sequence includes residues 1 to 1923 of the VIR-1 clone. The exact size of fragment lost from the PP1 genome has not been determined, but the Southern blotting analysis suggests that the deletion ends close to the breakpoint at base 1923. Thus the consensus sequence is encoded within the region which becomes deleted, and possibly marks the end of the deletion. In contrast, the location of the consensus sequence in the VIR-2 clone, in the PP2 genome, is before the fragment deleted in the transition from PP2 to CFA, and therefore probably marks the start of the deletion.

One possible explanation for the consensus sequence being present before the start of one deletion, and before the end of another is that the sequence is recognised by a cleaving enzyme. This enzyme is presumably influenced by other factors to determine which recognition sites should be cleaved in different environments, as the VIR-2 clone contains a second consensus sequence (upstream of base 1923), which is not cleaved in either rearrangement. In order to facilitate the two sequential deletions in the transition from PP1 to CFA a further copy of the

consensus sequence may be present within the unsequenced 800bp region, to mark the start of the PP1-PP2 deletion, and to be deleted itself in the PP2-CFA deletion.

The presence of internal copies of the repeat unit was also considered as a mechanism by which the C500 genome may recombine. The ladders observed, using the various clones as probe (see chapters 4 and 5), all produced large sized fragments, suggesting that if internal copies of the repeat units existed, there must be multiple copies present. In addition the WC11 mapping studies did not suggest the presence of internal repeats (Bridgen 1991).

In conclusion, the recombination mechanism which results in the C500 derivatives observed was not resolved in the course of this study, but a consensus sequence located close to the breakpoints was identified as a possible cleavage signal.

7.5 COMPARISON TO OTHER GAMMAHERPESVIRUSES

Infection of susceptible hosts with AHV-1 or OHV-2 results in MCF. The clinical features of the disease caused by these two agents are similar and the viruses are thought to be closely related (see chapter 1). The amplification of an OHV-2 sequence, identical in size and homologous to the equivalent AHV-1 fragment (see section 6.9), demonstrated further the similarities between the two viruses. In addition, this amplification also demonstrated that the ORF for peptide 1 is conserved between AHV-1 and OHV-2.

The most closely related completely sequenced gammaherpesvirus to AHV-1 is HVS. The two genomes share the common features of multiple copies of tandemly arrayed, terminally located repeat units flanking an internal unique sequence (Bornkamm *et al* 1976, Bridgen 1991). The 5' terminus of the HVS genome encodes the genes responsible for the transforming ability of HVS (Beisinger *et al* 1990, Jung *et al* 1991, Jung and Desrosiers 1991). The region of the C500 genome containing the DNA sequence which is rearranged as a result of *in vitro* propagation was arbitrarily labelled as being adjacent to the 3' end of the unique region in chapters 4 and 5. Preliminary observations from sequencing short regions throughout the C500 genome have indicated that, in general, collinearity exists between C500 and HVS (A Ensser, personal communication). If the rearranged

C500 region contains ORFs with a similar function to the 5' HVS sequences then the unique region of the C500 genome represented by VIR-1, VIR-2 and ATT-1 should be re-orientated such that sequence formerly discussed as being 3' terminal would become 5' terminal. In the remainder of this chapter the region originally described as 3' in the C500 genome will be referred to as 5' to facilitate easier comparisons to other γ-herpesviruses.

HVS strains have been shown to vary greatly in the 5' terminal 7kbp region (Medveczky *et al* 1984). A contiguous 4.5kbp sequence of this region was found to be unnecessary for viral replication in cultured cells, but essential for the oncogenic potential of the strain (Desrosiers *et al* 1984 and 1985). The genomic alterations occurring in the 5' region of C500 result in loss of virulence, but replication in cell culture is enhanced. Two attenuated strains derived from adjustments to culturing conditions, strain 11att and SMH1-VERO, both exhibited spontaneous deletions in the 7kbp 5' unique region. In strain 11att a terminal fragment of 2.3kbp was deleted, and in SMH1-VERO a deletion of approximately 6.6kbp occurred (Koomey *et al* 1984, Daniel *et al* 1980). The attenuated derivative of C500, in comparison, differed from the virulent parental virus by loss of >3.8kbp.

The *Herpesvirus ateles* (HVA) genome exhibits general collinearity to HVS, with the exception of the 5' 4kbp unique region. Infection of susceptible hosts with HVA results in similar characteristics to those observed in HVS-induced malignancy (Falk 1980), but none of the ORFs in the HVA terminal region demonstrate homology to the HVS transforming proteins. However, HVA does encode two U RNAs in this region, which share homology with the HSURs (DeGrand and Mulder 1992). The C500 DNA sequenced did not encode any U RNAs, however, the general loss of collinearity exhibited between HVS and HVA was reflected in the 5' unique region of C500.

The lymphocryptovirus EBV has also been completely sequenced. The organisation of the EBV genome is generally collinear to the HVS genome, however, the genes responsible for the transforming ability of EBV are not arranged collinearly to the STPs of HVS. In addition no nucleic acid or amino acid sequence homology exists between the transforming genes of HVS or EBV (Albrecht *et al* 1992). The ORFs identified in this study of AHV-1 did not demonstrate significant homology to the genes responsible for the transforming ability of HVS or EBV, despite the general

collinearity between HVS, (and by implication EBV), (A. Ensser, personal communication).

The genes involved in the transforming ability of EBV are also involved in latent expression of the EBV genome. The three best characterised latent EBV transcripts are: EBV nuclear antigen (EBNA) 1, which is involved in activating transcription of the latent genome; EBNA 2, which is required for the initiation of lymphocyte transformation, and latent membrane protein (LMP), which has oncogenic effects (reviewed by Miller 1990). AHV-1 may be expressed in a similar manner to EBV, thus the lymphoproliferative effects of AHV-1 observed in susceptible hosts may be the result of expression of products of latent rather than lytic expression. In the natural host initial infection with AHV-1 is associated with at least some productive replication and release of cell-free virus (Mushi and Rurangirwa 1981) which readily spreads by contagion. Persistance of AHV-1 infection in this host is then in a latent form. The course of infection in MCF susceptible hosts is obscure, but no contagious spread is observed and very limited virus transcription appears to occur, although cell-associated infectivity can be demonstrated. In affected tissues very little if any viral antigen or DNA can be demonstrated (see section 1.8) and it may be that virus transcription is limited to that associated with latency in the natural host, however, in the inappropriate host these products lead to the immune deregulation recognised as MCF. Alternatively expression of specific latent products may vary depending on the host. Studies of EBV-associated disease in immunocompromised individuals showed differential expression of latent products. In Burkitt's lymphoma tissue samples demonstrated EBNA 1 alone, while nasopharyngeal carcinoma tissues contained EBNA 1 and a proportion (40%) of tissues also demonstrated LMP. In contrast, tissues from lymphoproliferative disease associated with EBV contained EBNA 1, EBNA 2 and LMP (Cohen 1991).

In vitro propagation of AHV-1 resulted in attenuation, both of the C500 isolate and the WC11 isolate. In this artificial environment the result of AHV-1 infection of bovine monolayers is target cell lysis, suggesting that the virus is acting in a replicative rather than latent manner. Thus the genes encoding the peptides necessary for the establishment of latency would not be required *in vitro*. Perhaps these genes are deleted in the CFA genome, hence CFA virus would subsequently be unable to establish latency, and therefore virulence would be lost.

The gene encoding the latently expressed EBNA-2 protein is essential for the lymphocyte transformation associated with EBV (Hammershmidt and Sugden 1989). A laboratory derived strain, P3HR-1, which lacks a DNA segment that includes the EBNA-2 gene and part of the EBNA leader protein gene will not transform lymphocytes (King *et al* 1982, Rabson *et al* 1982). Recombinant transformation-competent EBV was produced by transfecting P3HR-1-infected cells with a cosmid vector containing EBNA-2 (Cohen *et al* 1989). Further studies suggested that EBNA-2 plays an important dual role in the process of B cell activation to the lymphoblastoid phenotype; the protein can have a direct role upon cellular gene expression and is also involved in activating the expression of a second virus encoded effector protein, LMP (Abbot *et al* 1990). The ORF for peptide 5, and the 800bp region, were deleted in the attenuated C500 genome, thus peptide 5 may have a similar role to that of EBNA-2 in the pathogenesis of MCF.

The switch between latency and replication of EBV is controlled by expression of a single polypeptide, designated ZEBRA (Bam HI fragment Z, EBV replication activator), (Grogan *et al* 1987). The possible AHV-1 equivalent(s) to this ZEBRA protein would be peptides 7, 8 and 9, as the sequence 5' to these ORFs is altered as the virus becomes more replication competent. This rearrangement may result in these peptides being located close to promoter sequences.

Certain defective EBVs with deleted and rearranged genomes do not establish latency (Miller *et al* 1984). The P3HR-1 strain of EBV was found to contain a proportion of defective genomes, termed het DNA. This het DNA was extensively deleted and rearranged, and the gene for ZEBRA was under positive regulation (reviewed by Miller 1990). The het DNA genome is an example of defective interfering particles (DIP). DIP are described as subgenomic deletion mutants, generated from infectious particles, which are capable of modulating infections. DIP frequently play an important part in viral persistance *in vitro*, and may, in some instances, modify viral infections *in vivo,* causing attenuation or persistance of an infection. Successive passages *in vitro*, using a high multiplicity of infection, frequently give rise to DIP (summarised by Roux *et al* 1990). However, DIP are not believed to be involved in the attenuation of AHV-1 observed as a result of *in vitro* propagation, due to the biological cloning of the virus preparations carried out at various time points during cell culture.

7.7 CONCLUSIONS

The aim of this project was to investigate the genomic arrangements occuring *in vitro* which result in loss of virulence of the C500 isolate of AHV-1. The variable regions observed initially from Sma I digests of the C500 derivatives were representative of two sequential genomic alterations which resulted in attenuation. The significance of these rearrangements was investigated by sequencing the three molecular clones derived from the C500 derivatives, and by assessing their location with regard to the terminal repeats of the molecule. The 3' region of the ORF which encodes peptide 5 was lost when the virus became attenuated, and therefore peptide 5 appeared to be the most likely candidate for involvement in the latency/transformation mechanism of the virus. In addition, the 800bp region adjacent to this ORF was also deleted in the transition from virulence to attenuation, therefore may encode peptide(s) involved in virulent infection of susceptible hosts. The failure to clone this region, and the inability to amplify the region using PCR suggested that the DNA encoded in this region contains obstructive structural features.

7.8 FUTURE WORK

The work carried out in this project offers many possibilities for future studies to extend the data presented. The 800bp uncloned and unsequenced deleted region requires to be cloned, using an enzyme other than Sma I. The previous attempts to clone and amlify this region suggest that sequencing of the region may be problematic, however, once cloned, this should be attempted.

The proposed genomic rearrangements of the C500 isolate during *in vitro* propagation need to be confirmed. One method to achieve this would be to carry out double digests of PP1 and PP2 DNA using various combinations of the restriction endonucleases Hind II, Hind III, Sac I and Sma I, as all fragment sizes could be predicted. This work was not undertaken in the present study due to difficulties in generating sufficient viral DNA.

The aim of this work was to determine the factors important for the virulence of the C500 isolate of AHV-1. The ORF for peptide 5 appears to be important in maintaining virulence, as loss of this ORF correlates with attenuation. Extensive

expression studies should be undertaken to establish whether this peptide is expressed *in vivo* and *in vitro*. Preliminary expression studies have consisted of Northern blot analysis of PP1 tissue culture mRNA products. No hybridisation to peptide 5 was observed, however, if peptide 5 is important for latency then it may not be expressed in tissue culture. Further studies to assess whether peptide 5 is expressed may require *in situ* hybridisation analysis of samples from affected animals. As the virus has only been detected in 1 in 10^4 cells in affected animals (Bridgen *et al* 1992), very little viral messsage may be present. The lymphoblastoid cell-lines developed from AHV-1 and OHV-2 affected animals (Wilkinson *et al* 1992, Reid *et al* 1983 and 1989a) may express peptide 5, therefore Northern blot analysis of mRNA produced by these cell-lines should be undertaken.

Studies to determine which of the nine hypothetical peptides identified in this study are expressed during *in vitro* propagation should also be undertaken. The proposed genomic rearrangements of the three C500 derivatives, coupled with the interpretation of the significance of the ORFs altered as described in this chapter suggest that one or all of peptides 7, 8 and 9 might be expressed in PP2 and CFA but not PP1 infection of cells of bovine lineage. In addition, expression of peptides 1-4 may be upregulated on infection with CFA virus.

Peptide 5 may be the AHV-1 analogue to the HVS transforming peptides (STPs), or to EBNA-2, and experiments to assess whether co-transfection of peptide 5 with the CFA genome may result in recombination events and virulent infection should be undertaken. Preliminary attempts to transfect bovine thyroid cells with the intact CFA genome have proved unsuccessful, however, to date the only method attempted has been by calcium chloride. Future studies should involve alternative methods of transfection such as electroporation. If peptide 5 is an analogue to the HVS STP peptides then it may be responsible for the generation of T lymphocytes immmortalised by AHV-1. HVS has been used to transform human $CD4^+$ and $CD8^+$ cells *in vitro* (Beisinger *et al* 1992). This immortalisation overcomes the problems of the necessity of periodic restimulation involved in the maintenance of human T cells. The immortalised cells can be used as a model to study signal transduction in activated human T lymphocytes (Broker *et al* 1993). Future work should include investigations into the possible use of AHV-1 to immortalise bovine T lymphocytes to facilitate immunological studies.

The sequence generated from the three C500 clones could be used to develop a diagnostic PCR. The observation that peptide 1 was also amplified from OHV-2 DNA with the C500 primers suggests that a PCR developed from this region of the clones would not be AHV-1 specific. PCR to determine whether peptide 5 is present in OHV-2 should also be undertaken, because if peptide 5 is vital to the pathogenic effects of MCF then a peptide 5 homologue would be anticipated in the OHV-2 genome.

The comparison of virulent and non-virulent AHV-1 derivatives described in this thesis has led to the identification of a variable region of the AHV-1 genome which is correlated with the ability of the virus to induce MCF when used to inoculate rabbits. The future applications of the data described may not only increase our understanding of the pathogenesis of MCF, but may also be used to explore the possibility of immunisation of susceptible animals with specific deletion mutants, and, more generally, as a tool in the study of bovine immunity.

References

Abbot, S.D., Rowe, M., Cadwaller, K., Ricksten, A., Gordon, J., Wang, F., Rymo, L. and Rickinson, A.B. (1990). Epstein-Barr virus nuclear antigen 2 induces expression of the virus encoded latent membrane protein. *Journal of Virology* **64,** 2126-2134.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1983). Molecular Biology of the Cell. Garland Publishing Inc. New York. p408

Albrecht, J.C. and Fleckenstein B. (1990). Structural organisation of the conserved gene block of *Herpesvirus saimiri* coding for DNA polymerase, glycoprotein B and major DNA binding protein. *Virology* **174,** 533-542.

Albrecht, J.C. and Fleckenstein B. (1992). Nucleotide sequence of HSUR 6 and HSUR 7, two small RNAs of *Herpesvirus saimiri. Nucleic Acids Research* **20**, 1810.

Albrecht, J.C., Nicholas, J., Biller, D., Cameron, K.R., Biesinger, B., Newman, C., Wittman, S., Craxton, M.A., Coleman, H., Fleckenstein, B. and Honess, R.W. (1992). Primary structure of the *Herpesvirus saimiri* genome. *Journal of Virology* **66,** 5047-5058.

Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Seguin, C., Tuffnell, P.S. and Barrell, B.G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* **310,** 207-211.

Bankier, A.T., Dietrich, W., Baer, R., Barrell, B.G., Colbere-Garapin, F., Fleckenstein, B. and Bodemer, W. (1985). Terminal repetitive sequences in *Herpesvirus Saimiri* virion DNA. *Journal of Virology* **55,** 133-139.

Baxter, S.I.F., Pow, I., Bridgen, A. and Reid, H.W. (1993). PCR detection of the sheep-associated agent of malignant catarrhal fever. *Archives of Virology* **132,** 145-159.

Biesinger, B., Muller-Fleckenstein, I., Simmer, B., Long, G., Wittman, S., Platzer, E., Desrosiers, R.C. and Fleckenstein, B. (1992). Stable growth of human T lymphocytes by *Herpesvirus saimiri. Proceedings of the National Academy of Sciences* USA **89,** 3116-3126.

Biesinger, B., Trimble, J.J., Desrosiers, R.C. and Fleckenstein, B. (1990). The divergence between two oncogenic *Herpesvirus saimiri* strains in a genomic region related to the transforming phenotype. *Virology* **176,** 505-514.

Birnboim, H.C. and Doly. J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7,** 1513-1523.

Biron, C.A., and Welsh, R.M. (1982). Activation and role of natural killer cells in virus infections. *Medical Microbiology and Immunology* **170,** 155-172.

Bornkamm, G.W., Delius, H., Fleckenstein, B., Werner, F.J. and Mulder, C. (1976). Structure of *Herpesvirus saimiri* genomes; arrangement of heavy and light sequences in the M genome. *Journal of Virology* **19,** 154-161.

Bornkamm, G. W., Delius, H., Zimber, U., Hudewentz, J. and Epstein, M. A. (1980). Comparison of Epstein-Barr virus strains of different origin by analysis of the viral DNAs. *Journal of Virology* **35,** 603-618.

Bridgen, A. (1991). The derivation of a restriction endonuclease map for *Alcelaphine Herpesvirus*-1 DNA. *Archives of Virology* **117,** 183-192.

Bridgen, A., Herring, A.J., Inglis, N.F. and Reid, H.W. (1989). Preliminary characterisation of the *Alcelaphine Herpesvirus*-1 genome. *Journal of General Virology* **70,** 1141-1150.

Bridgen, A., Munro, R. and Reid, H.W. (1992). The detection of *Alcelaphine Herpesvirus*-1 DNA by in situ hybridisation of tissues affected with malignant catarrhal fever. *Journal of Comparative Pathology* **106,** 351-359.

Bridgen, A. and Reid, H.W. (1991). Derivation of a DNA clone corresponding to the viral agent of sheep-associated malignant catarrhal fever. *Research in Veterinary Science* **50,**38-44.

Broker, B.M., Tsygankov, A.Y., Muller-Fleckenstein, I., Guse, A.H., Chitaev, N.A., Biesinger, B., Fleckenstein, B. and Emmrich, F. (1993). Immortalisation of human T cell clones by *Herpesvirus saimiri. Journal of Immunology* **151,** 1184-1192.

Bublot M., Lomonte, P., Lequarre, A.S., Albrecht, J.C., Nicholas, J., Fleckenstein, B., Pastoret, P-P. and Thiry, E. (1992). Genetic relationships between Bovine herpesvirus 4 and the gammaherpesviruses Epstein-Barr virus and *Herpesvirus saimiri. Virology* **190,** 654-665.

Bublot, M., Van Bressem, M.F., Thiry, E., Dubuisson, J. and Pastoret, P-P. (1990). *Bovine herpesvirus*-4 genome: cloning, mapping and strain variation analysis. *Journal of General Virology* **71,** 133-142.

Burrells, C. and Reid, H.W. (1991). Phenotypic analysis of lymphoblastoid cell lines derived from cattle and deer affected with "sheep-associated" malignant catarrhal fever. *Veterinary Immunology and Immunopathology* **29,** 151-161.

Buxton, D. and Reid, H. (1980). Transmission of malignant catarrhal fever to rabbits. *Veterinary Record* **106,** 243-245.

Buxton, D., Reid, H.W., Findlayson, J., Pow, I. (1984). Pathogenesis of "sheep-associated" malignant catarrhal fever in rabbits. *Research in Veterinary Science* **36,** 205-211.

Chou, P.Y. and Fasman, G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology* **47,** 45-148

Chung, C.T., Niemela, S.L. and Miller, R.H. (1989). One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Science USA* **86,** 2172-2175.

Cohen, J.I. (1991). Epstein-Barr virus gene expression in lymphoproliferative disease. *Leukaemia and Lymphoma* **3,** 235-240.

Cohen, J.I., Wang, F., Mannick, J. and Kieff, E. (1989). Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proceedings of the National Academy of Sciences* USA **86,** 9558-9562.

Cohrs, R. and Rouhandeh, H. (1987). Characterisation of the genome of *Herpesvirus sylvilagus. Intervirology* **28,** 181-184.

Coulson, A.F.W, Collins, J.F. and Lyall, A. (1987). Protein and nucleic acid database searching: a suitable case for parallel processing *The Computer Journal* **30,** 420-424.

Daniels, P.W., Sudarisman, and Purnomo Ronohardjo (Eds). Malignant Catarrhal Fever in Asian Livestock. Australian Centre for International Agricultural Research, Canberra 1988.

Daniel, M.D., Silva, D., Koomey, J.M., Mulder, C., Fleckenstein, B., Tamulerich, R., King, N.W., Hunt, R.D., Sehgal, P. and Falk, L. (1980). *Herpesvirus saimiri*: strain SMHI modification of oncogenicity p395-396. In D.S. Yohn, B.A. Lapin and J.R. Balkeslee (ed.) Advances in Comparative Leukaemia Research 1979. Elsevier/North Holland publishing Co. New York.

Daubney, R. and Hudson, J.R. (1936). Transmission experiments with bovine malignant catarrh. *Journal of Comparative Pathology* **49,** 63-89.

Dayhoff, M.O., Scharwtz, R.M. and Orcutt, B.C. (1978). In: Atlas of protein sequence and structure, volume 5, supplement 3 (Dayhoff, M.O. editor). National Biomedical Research Foundation, Washington D.C.

DeGrand, D.B. and Mulder, C. (1992). Sequence analysis of the right terminal region of L-DNA of Herpesvirus ateles. Abstract 73 from the 17th International Herpesvirus Workshop, Edinburgh.

Denholm, L.J. and Westbury, H.A. (1982). Malignant catarrhal fever in farmed Rusa deer (*Cervus timorensis*). 1.Clinico-pathological observations. *Australian Veterinary Journal* **58,** 81-87.

Desrosiers, R.C., Bakker, A., Kamine, J., Falk, L.A., Hunt, R.D. and King, N.W. (1985). A region of the *Herpesvirus saimiri* genome required for oncogenicity. *Science* **228,** 184-187.

Desrosiers, R.C., Burghoff, R.L., Bakker, A. and Kamine, J. (1984). Construction of replication competent *Herpesvirus saimiri* deletion mutants. *Journal of Virology* **49,** 343-348.

Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12,** 387-395.

Edington, N., Patel, J., Russell, P.H. and Plowright, W. (1979). The nature of the acute lymphoid proliferation in rabbits infected with the herpesvirus of bovine malignant catarrhal fever. *British Journal of Cancer* **15,** 1515-1522.

Edington, N. and Plowright, W. (1980). The protection of rabbits against the herpesvirus of malignant catarrhal fever by inactivated vaccines. *Research in Veterinary Science* **28,** 384-386.

Efstathiou, S., Ho, Y.M., Hall, S., Styles, C.J., Scott, S.D. and Gompels, U.A. (1990a). Murine herpesvirus 68 is genetically related to the gammaherpesviruses Epstein-Barr virus and *Herpesvirus saimiri. Journal of General Virology* **71,** 1365-1372.

Efstathiou, S., Ho Y.M., Minson, A.C. (1990b). Cloning and molecular characterisation of the murine herpesvirus 68 genome. *Journal of General Virology* **71,** 1355-1364.

Ehlers, B., Buhk, H.J. and Ludwig, H. (1985). Analysis of bovine cytomegalovirus genome structure: cloning and mapping of the monomeric polyrepetitive DNA unit, and comparison of European and American strains. *Journal of General Virology* **66,** 55-68.

Ensser, A. personal communication. Molekular biologische laboratorien, Schlosgarten, 4 D-8520, Erlangen, Germany.

Falk, L.A. (1980). Biology of *Herpesvirus saimiri* and *Herpesvirus ateles*.In "Oncogenic Herpesviruses" pp813-832. Edited by G. Klein. New York: Raven Press.

Feinberg, A.P and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6-13.

Fickett (1982) Recognition of protein coding regions in DNA sequences. *Nucleic Acids Research* **10,** 5303-5318.

Fleckenstein, B., Bornkamm, G.W. and Ludwig, H. (1975). Repetitive sequences in complete and defective genomes of *Herpesvirus saimiri*. *Journal of Virology* **15,** 398-406.

Fleckenstein, B., Bornkamm, G.W., Mulder, C., Werner, F-J., Daniel, M.D., Falk, L.A. and Delius, H. (1978). *Herpesvirus ateles* DNA and its homology with *Herpesvirus saimiri* nucleic acid. *Journal of Virology* **25,** 361-373.

Fleckenstein, B. and Mulder, C. (1980). Molecular biological aspects of *Herpesvirus saimiri* and *Herpesvirus ateles.* In "Viral Oncology" pp799-812. Edited by G. Klein. New York: Raven Press

Gilles, P.N. and Ogston, W. (1991). Structure and expression of the woodchuck herpesvirus genome. *Virology* **180,** 434-438.

Gompels, U.A., Craxton, M.A. and Honess, R.W. (1988). Conservation of gene organisation in the lymphotropic herpesviruses *Herpesvirus saimiri* and Epstein-Barr virus. *Journal of Virology* **62,** 757-767.

Green, I.R. and Sargan, D.R. (1991). Sequence of the cDNA encoding ovine tumor necrosis factor-α: problems with cloning by inverse PCR. *Gene* **109,** 203-210.

Grogan, E., Jenson, H., Countryman, J., Heston, L., Gradoville, L. and Miller G. (1987). Transfection of a rearranged viral DNA fragment, WZhet, stably converts latent Epstein-Barr viral infection to produce infection in lymphoid cells. *Proceedings of the National Academy of Sciences* **84,** 1332-1336.

Hammerschmidt, W. and Sugden, B. (1989). Genetic analysis of immortalising functions of Epstein-Barr virus in human lymphocytes. *Nature* **340,** 393-397.

Hatkin, J. (1980). Endemic malignant catarrhal fever at the San Diego Wild Life Animal Park. *Journal of Wildlife Diseases* **16,** 439-443.

Hayward, S.D., Nogee, L. and Hayward, G.S. (1980). Organisation of repeated regions within the Epstein-Barr virus DNA molecule. *Journal of Virology* **40,** 861-869.

Herring, A.J., Berrie, E., Reid, H.W. and Pow, I. (1983). Malignant Catarrhal Fever virus: purification and preliminary characterisation of the genome. Eighth International Herpesvirus Workshop p53.

Herring, A.J., Inglis. N.F., Ojeh, C.K., Snodgrass, D.R. and Menzies, J.D. (1982). Rapid diagnosis of rotavirus infection by direct detection of viral nucleic acid by silver-stained polyacrylamide gels. *Journal of Clinical Microbiology* **16,** 473-477.

Herring, A.J., Reid, H.W., Inglis, N. and Pow, I. (1989). Immunoblotting analysis of the reaction of wildebeest, sheep and cattle sera with the structural antigens of

Alcelaphine Herpesvirus-1 (malignant catarrhal fever virus). *Veterinary Microbiology* **19,** 205-215.

Heuschele, W.P., Fletcher, H.R., Oosterhuis, J., Jannsen, D. and Robinson, P.T. (1984). Epidemiologic aspects of malignant catarrhal fever in the USA. Proceedings of the 88th Annual Meeting of the U.S. Animal Health Association pp640-651.

Heuschele, W.P., Seal, B.S. and Kleiforth, R.B. (1988). Malignant catarrhal fever in wild ruminants. Internationalen Symposiums uber die Erkrankungen der Zoo- und Wildtiere. Sofia.

Honess. R.W. (1984). Herpes simplex and "The Herpes Complex": diverse observations and a unifying hypothesis. *Journal of General Virology* **65,** 2077-2107.

Honess, R.W., Gompels, U.A., Barrell, B.G., Craxton, W., Cameron, K.R., Staden, R., Chang, Y-N., and Hayward, G.S. (1989). Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNA may be diagnostic of differences in the states of their latent genomes. *Journal of General Virology* **70,** 837-855.

Hsu, D., Shih, L.M., Castro, A.E. and Zee, Y.C. (1990a). A diagnostic method to detect *Alcelaphine herpesvirus*-1 of malignant catarrhal fever using the polymerase chain reaction. *Archives of Virology* **114,** 259-263.

Hsu, D., Shih, L.M. and Zee, Y.C. (1990b). Nucleotide sequence of a 3.5kbp fragment of malignant catarrhal fever virus strain WC11. *Archives of Virology* **113,** 53-60.

Hunt, R.D. and Billups, L.H. (1979). Wildebeest-associated malignant catarrhal fever in Africa: a neoplastic disease of cattle caused by an oncogenic herpesvirus? *Comparative Immunology and Microbiology of Infectious Diseases* **2,** 275-283.

Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (1990). PCR Protocols: A Guide to Methods and Applications. Academic Press, Inc.

Jackson, I.J. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Research* **19,** 3795-3798.

Jacoby, R.O., Buxton, D. and Reid, H.W. (1988a). The pathology of wildebeest-associated malignant catarrhal fever in hamsters, rats and guinea-pigs. *Journal of Comparative Pathology* **98,** 99-109.

Jacoby, R.O., Reid, H.W., Buxton, D. and Pow, I. (1988b). Transmission of wildebeest-associated and sheep-associated malignant catarrhal fever to hamsters, rats and guinea-pigs. *Journal of Comparative Pathology* **98,** 91-98.

James, M.P., Neilson, F.J.A. and Stewart, W.J. (1975). An epizootic of malignant catarrhal fever. 1. Clinical and pathological observations. *New Zealand Veterinary Journal* **23,** 9-12.

Jameson, B.A. and Wolf, H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants. *Computer Applied Bioscience* **4,** 181-186.

Jung, J.U. and Desrosiers, R.C. (1991). Identification and characterisation of the *Herpesvirus saimiri* oncoprotein STP-C488. *Journal of Virology* **65,** 6953-6960.

Jung, J.U., Trimble, J.J., King, N.W., Beisinger, B., Fleckenstein, B.W. and Desrosiers, R.C. (1991). Identification of transforming genes of subgroup A and C strains of *Herpesvirus saimiri. Proceedings of the National Academy of Sciences USA* **88,** 7051-7055.

Kalunda, M., Dardiri, A.H. and Lee, K.M. (1981a). Malignant Catarrhal Fever. I: Response of American cattle to malignant catarrhal fever virus isolated in Kenya. *Canadian Journal of Comparative Medicine* **45,** 70-76.

Kalunda, M., Ferris, D.H., Dardiri, A.H. and Lee, K.M. (1981b). Malignant Catarrhal Fever III: Experimental infection of sheep, domestic rabbits and laboratory animals with malignant catarrhal fever virus. *Canadian Journal of Comparative Medicine* **45,** 310-314.

Katz, J., Seal, B. and Ridpath, J. (1991). Molecular diagnosis of Alcelaphine herpesvirus (malignant catarrhal fever) infections by nested amplification of viral DNA in bovine buffy coat specimens. *Journal of Veterinary Diagnostic Investigation* **3,** 193-198.

King, W., Dambaugh, T., Heller, M., Dowling, J. and Kieff, E. (1982). Epstein-Barr virus II. A variable region of the genome is included in the P3HR-1 deletion. *Journal of Virology* **43,** 979-986.

Klein, G. (1989). Viral latency and transformation: The strategy of Epstein-Barr virus. *Cell* **58,** 5-8.

Knust, E., Schirm, S., Dietrich, W., Bodemer, W., Kolb, E. and Fleckenstein, B. (1983). Cloning of *Herpesvirus saimiri* DNA fragments representing the entire L-region of the genome. *Gene* **25,** 281-289.

Koomey, J.M., Mulder, C., Burghoff, R.L., Fleckenstein, B. and Desrosiers, R.C. (1984). Deletion of DNA sequences in a nononcogenic variant of *Herpesvirus saimiri. Journal of Virology* **50,** 662-665.

Kyte, J. and Doolittle, R.K. (1981). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157,** 105-132.

Laimins, L.A., Khourg, G., Gorman, C., Haward, B. and Gruss, P. (1982). Host specific activation of transcription from tandem repeats from simian virus 40 and moloney murine sarcoma virus. *Proceedings of the National Academy of Sciences* USA **79,** 6453-6457.

Lee, S.I., Murthy, S.C.S., Trimble, J.J., Desrosiers, D.C. and Steitz, J.A. (1988). Four novel U RNAs are encoded by a herpesvirus. *Cell* **54,** 599-607.

Liggitt, H.D. and De Martini, J.C. (1980a). The pathomorphology of malignant catarrhal fever. I. Generalised lymphoid vasculitis. *Veterinary Pathology* **17,** 58-72.

Liggitt, H.D. and De Martini, J.C. (1980b). The pathomorphology of malignant catarrhal fever. II. Multisystemic epithelial lesions. *Veterinary Pathology* **17,** 73-83.

Liggitt, H.D., McChesney, A.F. and De Martini, J.C. (1980). Experimental transmission of bovine malignant catarrhal fever to a bison (*bison bison*). *Journal of Wildlife Diseases* **16,** 299-304.

Lipman, D.J., and Pearson, W.R. (1985). Rapid and sensitive protein similarity searches. *Science* **277,** 1435-1441.

Ludwig, H. (1983). Bovine herpesviruses. In 'The Herpesviruses', volume 2, p135-214. B. Roizman (ed). New York and London: Plenum Press.

McGeoch, D.J., Dalrymple, M.A., Davison, A.J., Dolan, A., Frame, M.C., McNab, D., Perry, L.J., Scott, J.E. and Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *Journal of General Virology* **69,** 1531-1574.

Maizel, J.V. and Lenk, R.P. (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proceedings of the National Academyof Sciences USA* **78,** 7665-7669.

Mandel, M. and Higa, A. (1970). Calcium dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53,** 159-162.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratories, New York.

Maniatis, T and Reed, R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325,** 673-677.

Mare, J.C. (1977). Malignant catarrhal fever, an emerging disease of cattle in the USA. *Proceedings of the 81st Annual Meeting of the US Animal Health Association* **81,** 151-157.

Medveczky, M.M., Geck, P., Clarke, C., Byrnes, J., Sullivan, J.L. and Medveczky, P.G. (1989a). Arrangement of repetitive sequences in the genome of *Herpesvirus Sylvaligus*. *Journal of Virology* **63,** 1010-1014.

Medveczky, P., Szomolanyi, E., Desrosiers, R.C. and Mulder. C. (1984), Classification of *Herpesvirus saimiri* into three groups based on extreme variation in a DNA region required for oncogenicity. *Journal of Virology* **52,** 938-944.

Medveczky, P., Szomolanyi, E., Hesselton, R., Degrand, D., Geck, P. and Medveczky, P.G. (1989b). *Herpesvirus saimiri* strains from three DNA subgroups have different oncogenic potentials in New Zealand white rabbits. *Journal of Virology* **63,** 3601-3611.

Melendez, L.V., Daniel, M.D. and Hunt, R.D. (1968). An apparently new herpesvirus from primary kidney cultures of the squirrel monkey (*saimiri sciureus*). *Laboratory Animal Care* **18,** 374-381.

Melendez, L.V., Hunt, R.D., Daniel, M.D., Garcia, E. and Fraser, C.E.O. (1969). *Herpesvirus saimiri* II: Experimentally induced malignant lymphoma in primates. *Laboratory Animal Care* **19,** 378-386.

Miller, G. (1984). Regions of the EBV genome involved in latency and lymphocyte immortalisation. *Progress in Medical Virology* **30,** 107-128.

Miller, G. (1990). The switch between latency and replication of Epstein-Barr virus. *Journal of Infectious Diseases* **161,** 833-844.

Miller, G., Rabson, M. and Heston, L. (1984). Epstein-Barr virus with heterogeneous DNA disrupts latency. *Journal of Virology* **50,** 174-182.

Murthy, S., Kamine, J. and Desrosiers, R.C. (1986). Viral-encoded small RNAs in *Herpesvirus saimiri* induced tumours. *EMBO Journal* **5,** 1625-1632.

Murthy, S.C.S., Trimble, J.J. and Desrosiers, R.C. (1989). Deletion mutants of *Herpesvirus saimiri* define an open reading frame necessary for transformation. *Journal of Virology* **63,** 3307-3314.

Mushi, E.Z., Jesset, D.M., Rurangirwa, F.R., Rossiter, P.B. and Karstad, L. (1981). Neutralising antibodies to malignant catarrhal fever herpesviruses in wildebeest nasal secretions. *Tropical Animal Health and Production* **13,** 55-56.

Mushi, E.Z. and Karstad, L. (1981). Prevalence of virus neutralising antibodies to malignant catarrhal fever virus in oryx (*Oryx beisa callotis*). *Journal of Wildlife Diseases* **17,** 467-470.

Mushi, E.Z., Karstad, L. and Jesset, D.M. (1980a). Isolation of bovine malignant catarrhal fever virus from ocular and nasal secretions of wildebeest calves. *Research in Veterinary Science* **29,** 168-171.

Mushi, E.Z. and Rurangirwa, F.R. (1981). Epidemiology of bovine malignant catarrhal fevers, a review. *Veterinary Research Communications* **5,** 127-142.

Mushi, E.Z., Rurangirwa, F.R. and Karstad, L. (1980b). Effect of levamisole on the course of malignant catarrhal fever virus infection in rabbits. *Tropical Animal Health Production* **13,** 112.

Oliver, R.E., Beatson, N.S., Cathcart, A. and Poole, W.S. (1983). Experimental transmission of malignant catarrhal fever to red deer (*Cervus elaphus*). *New Zealand Veterinary Journal* **31,** 209-212.

Osorio, F.A., Reed, D.E., Van der Maaten, M.J. and Metz, C.A. (1988). Comparison of the herpesviruses of cattle by DNA restriction endonuclease analysis and serological analysis. *American Journal of Veterinary Research* **46,** 2104-2109.

Needleman, S.B. and Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48,** 443-453.

Patel, J.R. and Edington, N. (1980). The detection of the herpesvirus of bovine malignant catarrhal fever in rabbit lymphocytes *in vivo* and *in vitro. Journal of General Virology* **48,** 437-444.

Patel, J.R. and Edington, N. (1981). The detection and behaviour of the herpesvirus of malignant catarrhal fever in bovine lymphocytes. *Archives of Virology* **68,** 321-326.

Pierson, R.E., Hamdy, F.M., Dardiri, A.H., Ferris, D.H. and Schloer, G.M. (1979). Comparison of African and American forms of malignant catarrhal fever; transmission and clinical signs. *American Journal of Veterinary Research* **40,** 1091-1095.

Pierson, R.E., Thake, D., McChesney, A.E. and Storz, J. (1973). An epizootic of malignant catarrhal fever in feedlot cattle. *Journal of the American Veterinary Medical Association* **163,** 349-350.

Plowright, W. (1965a). Malignant catarrhal fever in East Africa I- behaviour of the virus in free-living populations of blue wildebeest (*Gorgon taurinus taurinus*, Burchell). *Research in Veterinary Science* **6,** 56-58.

Plowright, W. (1965b). Malignant catarrhal fever in East Africa II- observations on wildebeest calves at the laboratory and contact transmission of the infection to cattle. *Research in Veterinary Science* **6,** 69-83.

Plowright, W. (1967). Malignant catarrhal fever in East Africa III - neutralising antibody in free-living wildebeest. *Research in Veterinary Science* **8,** 129-136.

Plowright, W. (1968). Malignant catarrhal fever. *Journal of the American Veterinary Medical Association* **152,** 795-804.

Plowright, W. (1986). Malignant catarrhal fever. *Revue Scientific et Technique Office International des Epizooties* **5,** 897-918.

Plowright, W., Ferris, R.D. and Scott, G.R. (1960). Blue wildebeest and the aetiological agent of bovine malignant catarrhal fever. *Nature* **188,** 1167-1169.

Plowright, W., Herniman, K.A.J., Jessett, D.M., Kalunda, M. and Rampton, C.S. (1975). Immunisation of cattle against the herpesvirus of malignant catarrhal fever; failure of inactivated vaccines with adjuvant. *Research in Veterinary Science* **19,** 159-166.

Plowright, W., Kalunda, M., Jesset, D.M. and Herniman, K.A.J. (1972). Congenital infection of cattle with the herpesvirus causing malignant catarrhal fever. *Research in Veterinary Science* **13,** 37-45.

Plowright, W., Macadam, R.F. and Armstrong, J.A. (1965). Growth and characterisation of the virus of bovine malignant catarrhal fever in East Africa. *Journal of General Microbiology* **39,** 253-266.

Rabson, M., Gradoville, L. Heston, L. and Miller, G. (1982). Non-immortalisiing P3J-HR-1 Epstein-Barr virus: a deletion mutant of its transforming Jijoye. *Journal of Virology* **44,** 834-844.

Reid, H.W. and Bridgen, A. (1991). Recovery of a herpesvirus from a roan antelope (*Hippotragus equinus*). *Veterinary Microbiology* **28,** 269-278.

Reid, H.W. and Buxton, D. (1984). Malignant catarrhal fever of deer. *Proceedings of the Royal Society of Edinburgh* **82B** 261-273.

Reid, H.W. and Buxton, D. (1985). Immunity and pathogenesis of malignant catarrhal fever. In Immunity to Herpesvirus Infections of Domestic Animals. Eds. P.-P. Pastoret, E Thiry and J Saliki. Report EUR9737 En, Commission of the European Communities.

Reid, H.W., Buxton, D., Berrie, E., Pow, I. and Finlayson, J. (1984). Malignant catarrhal fever. *Veterinary Record* **114,** 581-583.

Reid, H.W., Buxton, D., Berrie, E., Pow, I. and Findlayson, J. (1985). Culture of malignant catarrhal fever agent. Biology of Deer Production. *The Royal Society of New Zealand Bulletin* **22,** 143-146.

Reid, H.W., Buxton, D., Corrigal, W., Hunter, A.R., McMartin, D.A. and Rushton, R. (1979). An outbreak of malignant catarrhal fever in red deer (*Cervus elaphus*). *Veterinary Record* **104,** 120-123.

Reid, H.W., Buxton, D., Pow, I. and Findlayson, J. (1989a). Isolation and characterisation of lymphoblastoid cells from cattle and deer affected with "sheep-associated" malignant catarrhal fever. *Research in Veterinary Science* **47,** 90-96.

Reid, H.W., Buxton. D., Pow, I., Findlayson, J. and Berrie, E.L. (1983). A cytotoxic T-lymphocyte line propagated from a rabbit infected with sheep-associated malignant catarrhal fever. *Research in Veterinary Science* **34,** 109-113.

Reid, H.W., Jacoby, R.O., Pow, I. and Buxton, D. (1984). Culture of large granular lymphocytes from rats with a herpesvirus-associated lymphoproliferative condition. *Immunobiology* **167,** 192.

Reid, H.W., Plowright, W. and Rowe, L.W. (1975). Neutralising antibody to herpesvirus derived from wildebeest and hartebeest in wild animals in East Africa. *Research in Veterinary Science* **18,** 269-273.

Reid, H.W., Pow, I. and Buxton, D. (1989b). Antibody to *Alcelaphine herpesvirus*-1 (AHV-1) in hamsters experimentally infected with AHV-1 and the "sheep-associated" agent of malignant catarrhal fever. *Research in Veterinary Science* **47,** 383-386.

Reid, H.W. and Rowe, L. (1973). The attenuation of a herpesvirus (malignant catarrhal fever virus) isolated from Hartebeest (*Alcelaphus buselaphus cokei* Gunther). *Research in Veterinary Science* **15,** 144-146.

Roizman, B. (1982). The *Herpesviridae*: general description, taxonomy and classification. In "The Herpesviruses" Vol. 1. pp1-23. Edited by B. Roizman. New York and London: Plenum Press.

Roizman, B. (1990). *Herpesviridae*: A brief Introduction. *Virology*, Second Edition, edited by B.N. Fields, D.M. Knipe et al. Raven Press Ltd., New York.

Roizman, B., Carmichael, L.E., Deinhart, F., de-The, G., Nahmias, A.J., Plowright, W., Rapp, F., Sheldrick, P., Takahashi, M. and Wolf, K. (1981). *Herpesviridae.* Definition, provisional nomenclature, and taxonomy. *Intervirology* **16,** 201-217.

Roizman, B., Desrosiers, D.C., Fleckenstein, B., Lopez, C., Minson, A.C. and Studdert, M.J. (1992). The family *Herpesviridae*: an update. *Archives of Virology* **123,** 425-449.

Rossiter, P.B. (1981a). Antibodies to malignant catarrhal fever virus in sheep sera. *Journal of Comparptive Pathology* **91,** 303-311.

Rossiter, P.B. (1981b). Immunofluorescence and immunoperoxidase techniques for detecting antibodies to malignant catarrhal fever in infected cattle. *Tropical Animal Health Production* **13,** 189-192.

Rossiter, P.B. (1982a). Attempts to protect rabbits against challenge with virulent, cell-associated, malignant catarrhal fever virus. *Veterinary Microbiology* **7,** 419-425.

Rossiter, P.B. (1982b). Immunoglobulin response of rabbits infected with malignant catarrhal fever virus. *Research in Veterinary Science* **33,** 120-122.

Rossiter, P.B. and Jessett, D.M. (1980). A complement fixation test for antigens of and antibodies to malignant catarrhal fever virus. *Research in Veterinary Science* **28,** 228-233.

Rossiter, P.B., Jessett, D.M., Mushi, E.Z., and Karstad, L. (1980). Antibodies to malignant catarrhal fever virus antigens in the sera of normal and naturally infected cattle in Kenya. *Research in Veterinary Science* **29,** 235-239.

Rossiter, P.B., Mushi, E.Z. and Plowright, W. (1977). The development of antibodies in rabbits and cattle infected experimentally with an African strain of malignant catarrhal fever virus. *Veterinary Microbiology* **2,** *57-66*

Rossiter, P.B., Mushi, E.Z. and Plowright, W. (1978). Antibody response in cattle to early antigens of malignant catarrhal fever virus. *Research in Veterinary Science* **25,** 207-210.

Roux, L., Simon, A.E. and Holland, J.J. (1991). Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo. Advances in Virus Research* **40,** 181-211.

Russell, G.H. (1980). Malignant catarrhal fever virus in rabbits reproduction of clinical disease by cell-free virus and partial protection against such disease by vaccination with inactivated virus. *Veterinary Microbiology* **5,** 161-163.

Ruth, G.R., Reed, D.E., Daley, C.A., Vorhies, M.W., Wohlgemuth, K. and Shave, H. (1977). Malignant catarrhal fever in bison. *Journal of the American Veterinary Medical Association* **171,** 913-917.

Rweyemamu, M.M., Karstad, L., Mushi, E.Z., Otema, J.C., Jesset, D.M., Rowe, L., Drevemo, S. and Grootenhuis, J.G. (1974). Malignant catarrhal fever virus in nasal secretions of wildebeest: a probable mechanism for virus transmission. *Journal of Wildlife Diseases* **10,** 478-486.

Rweyemamu, M.M., Mushi, E.Z., Rowe, L. and Karstad, L. (1976). Persistant infection of cattle with the herpesvirus of malignant catarrhal fever and observations on the pathogenesis of the disease. *British Veterinary Journal* **132,** 393-400.

Saiki, R.F., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239,** 487-491.

Saiki, R.F., Scharf, S., Faloona, F., Miller, K.B., Horn, G.T., Erlich, H.A. and Arnhein, N. (1985). Enzymatic amplification of beta-globin genomic DNA sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **230,** 1350-1354.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: a laboratory manual, 2nd edition. Cold Spring Harbor Laboratory, New York.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74,** 5463-5467.

Schirm, S., Muller, I., Desrosiers, D.C. and Fleckenstein, B. (1984). HVS DNA ina lymphoid cell-line established by *in vitro* transformation. *Journal of Virology* **49,** 938-946.

Seal, B.S., Kleiforth, R.B., Castro, A.E. and Heuschele, W.P. (1988). Replication of Alcelaphine herpesviruses in various cell culture systems and subsequent purification of the virus. *Journal of Tissue Culture Methods* **11,** 49-56.

Seal, B.S., Klieforth, R.B. and Heuschele, W.P. (1987). Evidence for variation among malignant catarrhal fever virus isolates. *Proceedings of the 91st Annual Meeting of the U.S. Animal Health Association* 527-542.

Seal, B.S., Klieforth, R.B., Welch, W.H. and Heuschele, W.P. (1989). Alcelaphine herpesviruses 1 and 2. SDS-PAGE analysis of virion polypeptides, restriction endonuclease analysis of genomic DNA and virus replication restriction in different cell types. *Archives of Virology* **106,** 301-320.

Selman, I.E., Wiseman, A., Murray, M. and Wright, N.G. (1974). A clinico-pathological study of bovine malignant catarrhal fever in Great Britain. *Veterinary Record* **94,** 483-490.

Selman, I.E., Wiseman, A., Wright, N.G. and Murray, N. (1978). Transmission studies with bovine malignant catarrhal fever. *Veterinary Record* **102,** 252-257.

Shih, L-M, Zee, Y.C. and Castro, A.E. (1989). Comparison of genomes of malignant catarrhal fever-associated herpesviruses by restriction endonuclease analysis. *Archives of Virology* **109,** 145-151.

Short, J.M.. Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988). λ ZAP: bacteriophage L expression vector with *in vivo* excision properties. *Nucleic Acids Research* **16,** 7583-7600.

Smith, G.E. and Summers, M.D. (1980). The bidirectional transfer of DNA and RNA to nitrocellulose of diazobenzyloxymethyl-paper. *Analytical Biochemistry* **109,** 123-129.

Smith, T.F. and Waterman, M.S. (1981). Identification of common molecular sequences. *Journal of Molecular Biology* **147,** 195-197.

Stamminger, T., Honess, R.W., Young, D.F., Bodemer, W., Blair, E.D. and Fleckenstein, B. (1987). Organisation of terminal reiterations in the virion DNA of *Herpesvirus saimiri. Journal of General Virology* **68,** 1049-1066.

Van Santen, V.L. (1993). Characterisation of a bovine herpesvirus 4 Immediate-early RNA encoding a homolog of the Epstein-Barr virus R transactivator. *Journal of Virology* **67,** 773-784.

Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy ofSciences USA* **76,** 615-619.

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A. and Weiner, A.M. (1987). Molecular Biology of the Gene (fourth edition). Benjamin/Cummings Publishing Co. Inc

Westbury, H.A. and Denholm, L.J. (1982). Malignant catarrhal fever in farmed Rusa deer (Cervus timorensis) 2. Animal transmission andvirological studies. *Australian Veterinary Journal* **58,** 88-92.

Whitenack, D.L., Castro, A.E. and Kocan, A.A. (1981). Experimental malignant catarrhal fever (African form) in white-tailed deer. *Journal of Wildlife Diseases* **17,** 443-451.

Wilkinson, J.M., Galea-Lauri, J. and Reid H.W. (1992). A cytotoxic rabbit T-cell line infected with a γ-herpesvirus which expressed CD8 and class II antigens. *Immunology* **77,** 106-108.

Yannisch-Perron, C., Viera, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors *Gene* **33,** 103-119.

Appendices

Appendix 1 Suppliers

AMERSHAM - Amersham, Buckinghamshire, U.K.
BDH - Poole, Dorset, U.K.
BECKMAN - High Wickham, Buckinghamshire, U.K.
BIORAD - Hemel Hempstead, Hertfordshire, U.K.
BOEHRINGER MANNHEIM - Lewes, East Sussex U.K.
CAMBIO LIMITED - Cambridge, U.K.
DIFCO - East Molesey, Surrey, U.K.
FUNGAZONE - E.R. Squibb, Hounslow, Middlesex, U.K.
FISONS (SCOTLAB) - Coatbridge, U.K.
GENETIC RESEARCH INSTRUMENTATION LIMITED - Dunmow, Essex, U.K.
GIBCO BRL - Uxbridge, Middlesex, U.K.
HYBAID - Teddington, Middlesex, U.K.
NOVABIOCHEM - Nottingham, U.K.
OSWELL DNA - Edinburgh, U.K.
PERKIN ELMER CETUS - Norwalk, Connecticut, U.S.A.
PHARMACIA - Milton Keynes, U.K.
PROMEGA - Southampton, U.K.
SIGMA - Poole, Dorset, U.K.
STRATAGENE - Cambridge, U.K.
STRATECH SCIENTIFIC LIMITED - London, U.K.

Appendix 2 Description of the experiments carried out to map the ATT-1, VIR-1 and VIR-2 clones

A2.1 The mapping of ATT-1

The ATT-1 clone was initially mapped with the restriction endonucleases (REs) which have sites in the polylinker of pBS+ (see figure 2.2). The REs which had sites in the insert are listed in figure A2.1a, with the relevant sizes calculated accurately following the complete sequencing of the clone as described in chapter 3. The location of each RE site in the insert relative to its location in the polylinker allowed mapping if a site was present only once in the insert, as only two fragments would be produced on digestion of ATT-1, and a process of simple deduction located some RE sites. Thus the insert was mapped for Hind II, Hind III and Sph I on RE data alone. In addition the insert fragments ligated to pBS+ were also mapped for Kpn I and Pst I. This initial mapping data is shown in figure A2.1b.

The clone was mapped for Kpn I by using the 2572 and 339bp fragments respectively to probe Hind II and Hind III digests of ATT-1. The 2572bp fragment hybridised to both Hind II and both Hind III fragments, while the 339bp fragment hybridised to the Hind II 2857bp and the Hind III 3197bp fragments respectively. This meant that the 2572bp Kpn I fragment was located adjacent to the 5' end of the ATT-1 clone.

The clone was mapped for Pst I by probing Hind II, Hind III, Sph I and Kpn I digests with the Pst I 165, 216, 342 and 2039bp fragments respectively. The 165bp fragment hybridised to the Sph I 792bp fragment, the Kpn I 921bp fragment, the Hind II 2857bp fragment and the Hind III 3197bp fragment, implying that this fragment was located close to the 5'-end of the insert. The Pst I 216 and 342bp fragments both hybridised to the Sph I 6244bp fragment, the Kpn I 2572bp fragment, the Hind II 2851bp fragment and the Hind III 3197bp fragment, which meant that neither fragment was adjacent to the 5'-end of the clone. Size constraints meant that the 165bp fragment must be adjacent to the 5'-end of the clone, with the 2039bp fragment next. This was reinforced by the fact that the 2039bp Pst I fragment hybridised to both Sph I fragments, all three Kpn I fragments and the smaller Hind II and Hind III fragments. The orientation of the 216 and the 342 fragments was resolved by analysis of partial digests of the ATT-1 clone. The

Figure A2.1

Initial mapping data for ATT-1

Figure A2.1a ATT-1 RE fragment sizes

RE	Fragment sizes (in bp)
Hind II	4179, 2857
Hind III	3839, 3197
Kpn I	4125, 2572, 339
Pst I	4274, 2039, 342, 216, 165
Pvu II	2793, 1761, 1638, 844
Sph I	6244, 792

(total insert size 3832bp, plasmid size 3204bp)

Figure A2.1b Map of ATT-1 based on the restriction data in figure A2.1a

Key

H - Hind III	L - Sph I
I - Hind II	P - Pst I
K - Kpn I	S - Sma I

clone was linearised with Hind III, then Pst I was added at a final dilution of 0.2 units per μg of DNA. Aliquots of this digest were removed at 5, 10, 20 and 40 minute intervals, and the enzyme was inactivated immediately by incubation at 60°C. The resultant digests were then resolved by electrophoresis using 0.6% agarose, transferred to nylon filters and probed with DIG-labelled pBS+. The differences in the size of the bands and their order represented the size of the Pst I fragments and the order of these in the plasmid.

The Pvu II sites in the ATT-1 clone were mapped by using the 844, 1638 and 1761bp fragments respectively to probe Pst I digests. Pvu II does not have a site in the polylinker of pBS+, however it does have sites to either side of the polylinker hence the fragments to the 5' and 3' ends of the insert include pBS DNA. The 844bp fragment recognised Pst I fragments of 2039 and 342bp, suggesting a central location, while the 1638bp fragment recognised the Pst I 4274, 342 and 216bp fragments, locating this Pvu II fragment at the 3' end of the clone. Finally the 1761bp fragment recognised the 2039 and 165bp Pst I fragments implying a 5' location for this fragment.

The completed map of ATT-1 is presented in figure 3.6.

A2.2 The mapping of VIR-1

The VIR-1 clone was initially mapped following the method described for the ATT-1 clone. The REs which had sites in the insert and the relevant fragment sizes are exhibited in figure A2.2a, with a map of this data in figure A2.2b.

The Pst I map of the VIR-1 clone was completed first as the large number of Pst I fragments would facilitate accurate mapping of the remaining REs. The Pst I mapping was achieved in an identical manner to that described for ATT-1, however the plasmid was first linearised with Eco RI in place of Hind III.

The Sac I sites were located by using the 167 and the 2078bp fragments to probe Pst I digests. The 167bp fragment recognised Pst I fragments of 1139, 345 and 94bp, therefore this fragment was judged to be adjacent to the 5' Sac I fragment, with the 1139bp fragment located at the 3'-end of the insert. This was reinforced by the Pst I fragments that the 1139bp Sac I fragment hybridised to. Similarly the Hind II 340 and 1686bp fragments were used to probe Pst I digests, with the 340bp fragment hybridising to the Pst I 3' terminal 435bp fragment, thus suggesting

Figure A2.2

Initial mapping data for VIR-1

Figure A2.2a VIR-1 RE fragment sizes

RE	Fragment sizes (in bp)
Hind II	4717, 1686, 340
Pst I	4137, 1139, 435, 345, 342, 216, 84, 45
Pvu II	2793, 2061, 1036, 844
Sac I	4498, 2078, 167
Sph I	3411, 3332

(total insert size 3539bp, plasmid size 3204bp)

Figure A2.2b Map of VIR-1 based on the restriction data in figure A2.2a

Key

C - Sac I
H - Hind III
I - Hind II
P - Sph I
L - Pst I

terminal location, and the 1686bp Hind II fragment was shown to be the centrally located Hind II fragment.

The Pvu II sites were mapped by using the 844, 1036 and 2061bp fragments to probe Pst I digests. The 844bp fragment hybridised to Pst I fragments of 1139 and 342bp, and therefore was the centrally located Pvu II fragment. The 1036bp fragment hybridised to the Pst I 435, 342 and 216bp fragments, thus was located at the 3'-end of the clone, whilst the 2016bp fragment was found to be located at the 5'-end by hybridisation.

The completed map of VIR-1 is presented in figure 3.7.

A2.3 The mapping of VIR-2

The VIR-2 clone was initially mapped following the method described for ATT-1, using the REs in the polylinker of pBS SK+. REs which had sites in the insert and the relevant fragment sizes are exhibited in figure A2.3a, with a map of this data in figure A2.3b. Sph I was included in the mapping of VIR-2 although it does not have a site in the polylinker of pBS SK+.

The clone was mapped for Pst I by analysis of partial digests as described for ATT-1, with the plasmid first linearised with Eco RI instead of Hind III.

The insert was mapped for Kpn I using the 339 and 1937bp fragments to probe Pst I digests. The 339bp fragment hybridised to the Pst I fragment of 2039bp only, while the 1937bp fragment hybridised to Pst I fragments of 2039, 435, 342 and 216bp. This data meant that the 339bp Kpn I fragment was located centrally, and the 1937bp fragment was located at the 3'-end of the insert.

The insert was mapped for Sph I by double-digesting VIR-2 with Sph I and and Eco RI (a polylinker enzyme with no sites in the insert). This enabled location of the Sph I at 1123bp from the 5'-end of the clone.

The plasmid was mapped for Pvu II by using the 844, 998 and 2131bp fragments to probe Pst I digests of VIR-2. The 844bp fragment recognised Pst I fragments of 2039 and 342bp, suggesting a central location. The Pvu II 998bp fragment recognised Pst I fragments of 435, 342 and 216bp, implying a 3' location, while the 2131bp fragment recognised Pst I fragments of 2039 and 496bp, thus was

Figure A2.3

Initial mapping data for VIR-2

Figure A2.3a VIR-2 RE fragment sizes

RE	Fragment sizes (in bp)
Hind II	3298, 3188
Kpn I	4895, 1252, 339
Pst I	3393, 2039, 496, 342, 216
Pvu II	2513, 2131, 998, 844
Sph I	6486

(total insert size 3528bp, plasmid size 2958bp)

Figure A2.1b Map of ATT-1 based on the restriction data in figure A2.1a

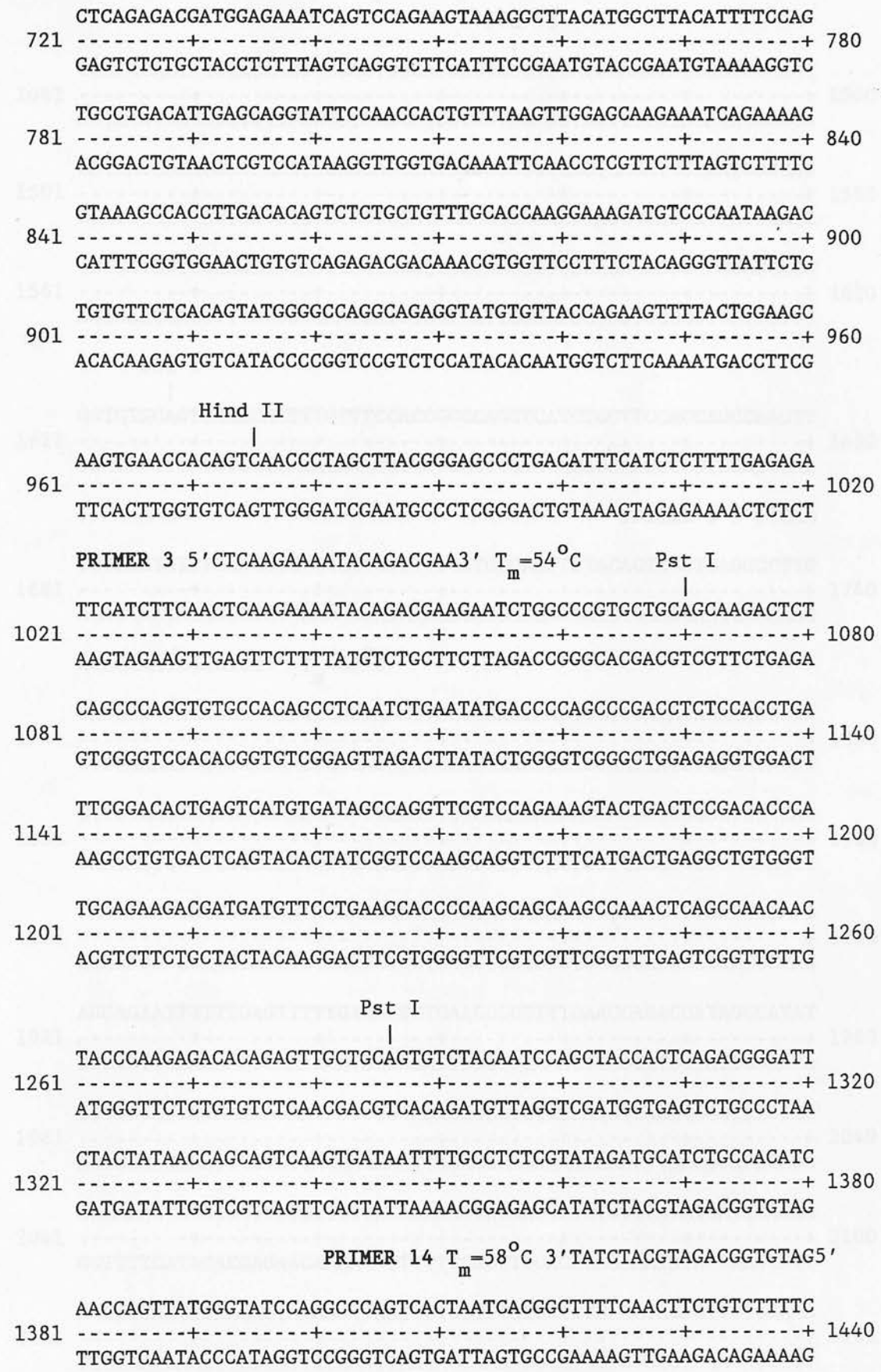
Key

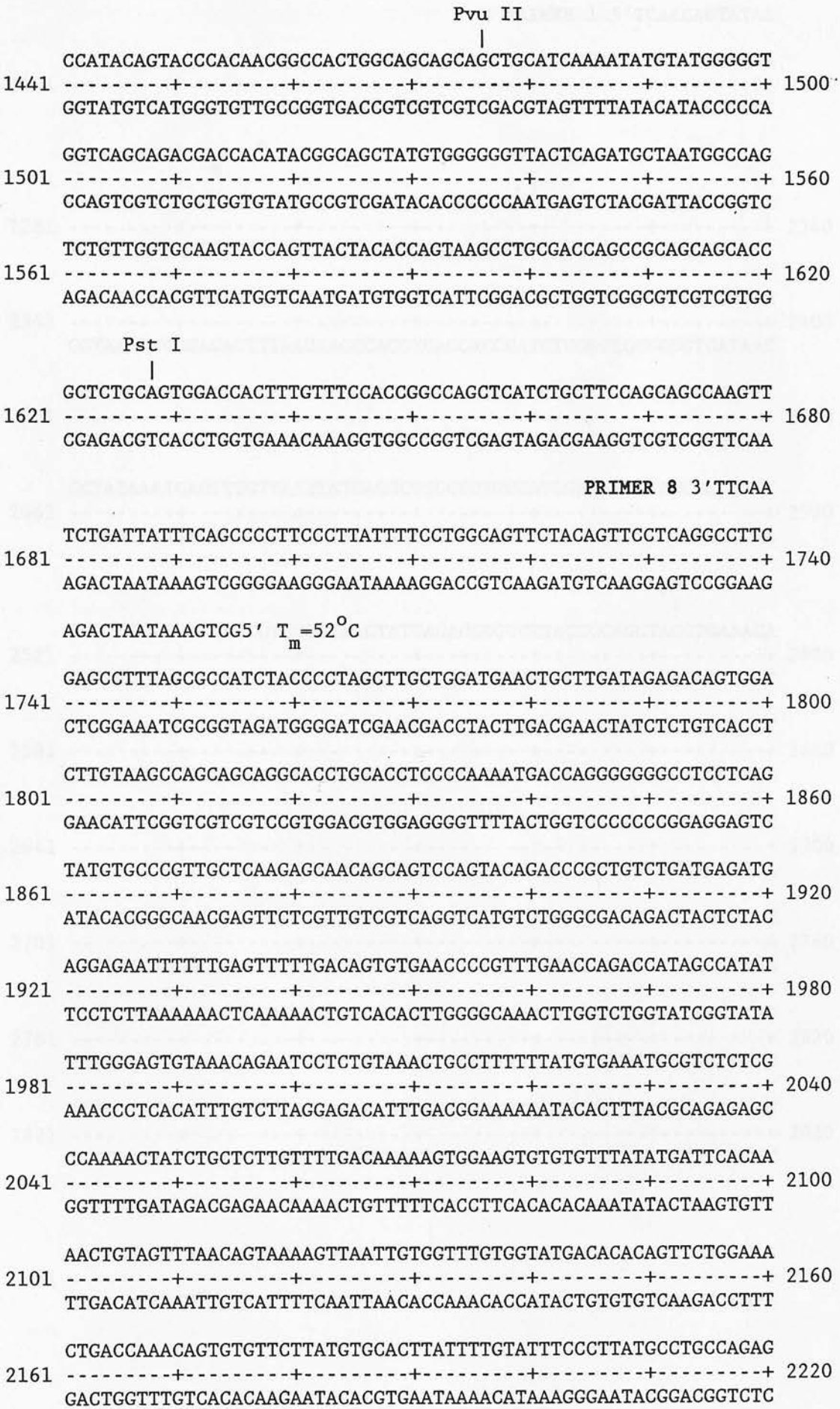
H - Hind III
I - Hind II
K - Kpn I
P - Pst I

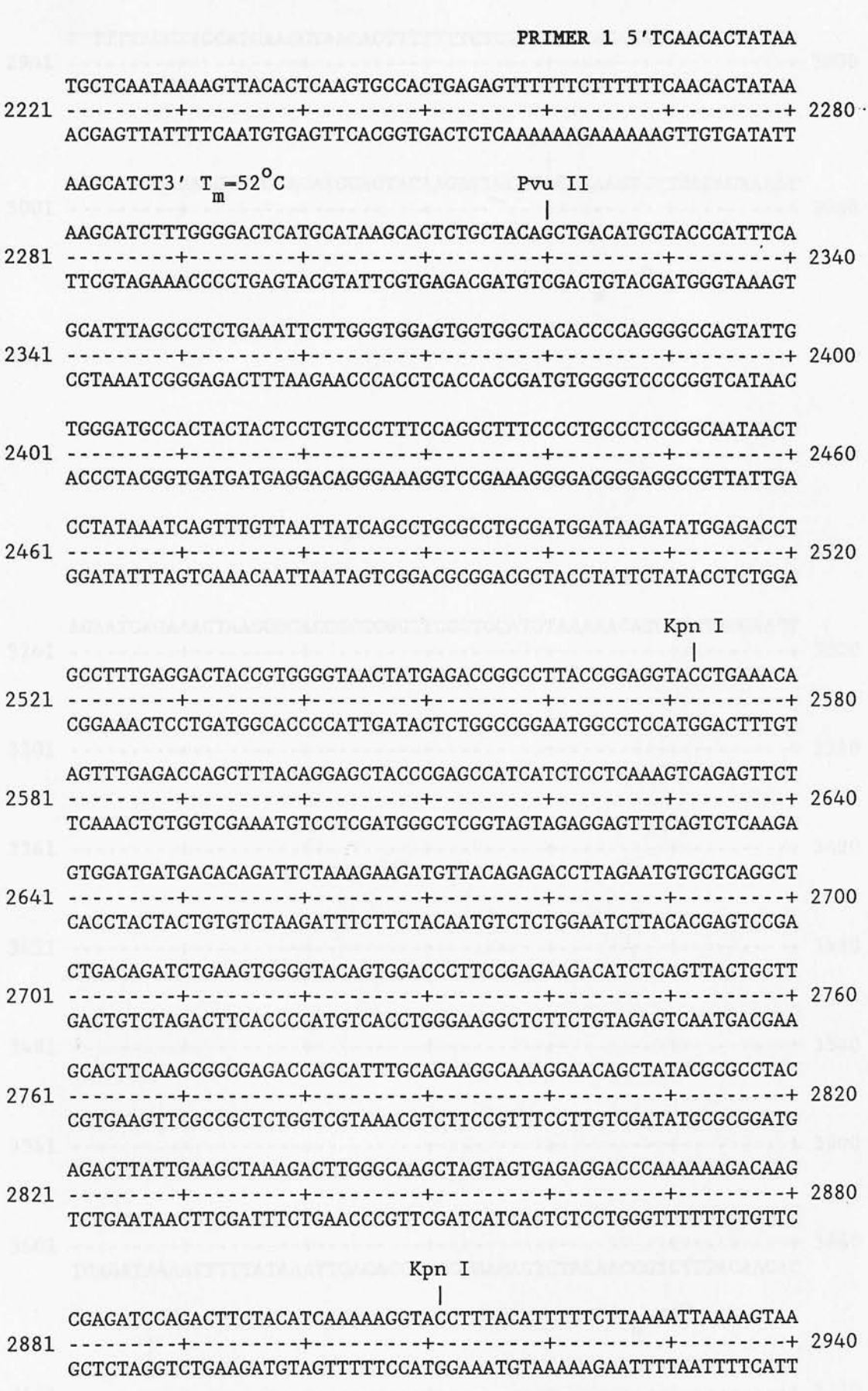
located at the 5'-end of the insert. The completed map of VIR-2 is presented in figure 3.11.

Appendix 3.1 The complete sequence of ATT-1

```
     GGGAAACAGTAAACCACATGTGTTGGACTTAAAGTTTGCTTGATCTAATTACTTGATTAT
   1 ---------+---------+---------+---------+---------+---------+ 60
     CCCTTTGTCATTTGGTGTACACAACCTGAATTTCAAACGAACTAGATTAATGAACTAATA

      PRIMER 12 3'GGTGTACACAACCTGAATTT5' Tm=56°C

     GTAACAGCCTTTTATACTAACATATTTATTTTTAGGATGATTAATAATAAACAATGACTA
  61 ---------+---------+---------+---------+---------+---------+ 120
     CATTGTCGGAAAATATGATTGTATAAATAAAAATCCTACTAATTATTATTTGTTACTGAT

     TCCTAAAAATGTTTTTGCATACTTTAAGGTTTAAAAATCTAAATTATATTTTTTAAATCA
 121 ---------+---------+---------+---------+---------+---------+ 180
     AGGATTTTTACAAAAACGTATGAAATTCCAAATTTTTAGATTTAATATAAAAAATTTAGT

     TTTTACAGAGGAGACTATCAAGACCAATATGCATAGATGACTTCATAGACATCACTGCTG
 181 ---------+---------+---------+---------+---------+---------+ 240
     AAAATGTCTCCTCTGATAGTTCTGGTTATACGTATCTACTGAAGTATCTGTAGTGACGAC

     ATCTTGGTGATACAATAGGGGCTGCTTTAAAGTCCTTCCAGCAAAATAATGCTTGCACTC
 241 ---------+---------+---------+---------+---------+---------+ 300
     TAGAACCACTATGTTATCCCCGACGAAATTTCAGGAAGGTCGTTTTATTACGAACGTGAG

     AAGAACAGAGTGAGCAGTTTACCAGGGAGTATATGATGTCTGCAAAATATTCTCCAGGAG
 301 ---------+---------+---------+---------+---------+---------+ 360
     TTCTTGTCTCACTCGTCAAATGGTCCCTCATATACTACAGACGTTTTATAAGAGGTCCTC

     AACAAATTTAGAAATGAAATGTTTGGATTTGTTGCTGATATGAACCTGCTAAATCTATTT
 361 ---------+---------+---------+---------+---------+---------+ 420
     TTGTTTAAATCTTTACTTTACAAACCTAAACAACGACTATACTTGGACGATTTAGATAAA

     GCCCTGTTCAGAAGCTATAAGCAGAGAGTTAGGACCACTTTGGAAAACAGCTTCTATGTG
 421 ---------+---------+---------+---------+---------+---------+ 480
     CGGGACAAGTCTTCGATATTCGTCTCTCAATCCTGGTGAAACCTTTTGTCGAAGATACAC

     CCACTGCCTCATCTCAAATATTAGATTCTTTTTAGAAAGAGTTATAAGGCTCTCTGACAA
 481 ---------+---------+---------+---------+---------+---------+ 540
     GGTGACGGAGTAGAGTTTATAATCTAAGAAAAATCTTTCTCAATATTCCGAGAGACTGTT

                          PRIMER 11 5'CATTCTCCCACAAGAGCTGG3' Tm=62°C

     ATGGTTTTTGTTAGCCCCCTGCAATGGCCTCATTCTCCCACAAGAGCTGGCGAAGGAAAT
 541 ---------+---------+---------+---------+---------+---------+ 600
     TACCAAAAACAATCGGGGGACGTTACCGGAGTAAGAGGGTGTTCTCGACCGCTTCCTTTA

                                  Hind III
                                     |
     GTATGTTCTCCTGTCCGAAGCCAGGGGGAAAGCTAAGCTTTGAACCAAGGGAGAATGTTT
 601 ---------+---------+---------+---------+---------+---------+ 660
     CATACAAGAGGACAGGCTTCGGTCCCCCTTTCGATTCGAAACTTGGTTCCCTCTTACAAA

     AGCGGAGGGAGACAAAACATGATGAATGCCGCCAAGAAGGTTTTGACAGTGTACTCTAGC
 661 ---------+---------+---------+---------+---------+---------+ 720
     TCGCCTCCCTCTGTTTTGTACTACTTACGGCGGTTCTTCCAAAACTGTCACATGAGATCG
```


CTCAGAGACGATGGAGAAATCAGTCCAGAAGTAAAGGCTTACATGGCTTACATTTTCCAG
721
GAGTCTCTGCTACCTCTTTAGTCAGGTCTTCATTTCCGAATGTACCGAATGTAAAAGGTC
TGCCTGACATTGAGCAGGTATTCCAACCACTGTTTAAGTTGGAGCAAGAAATCAGAAAAG
781
ACGGACTGTAACTCGTCCATAAGGTTGGTGACAAATTCAACCTCGTTCTTTAGTCTTTTC
GTAAAGCCACCTTGACACAGTCTCTGCTGTTTGCACCAAGGAAAGATGTCCCAATAAGAC
841
CATTTCGGTGGAACTGTGTCAGAGACGACAAACGTGGTTCCTTTCTACAGGGTTATTCTG
TGTGTTCTCACAGTATGGGGCCAGGCAGAGGTATGTGTTACCAGAAGTTTTACTGGAAGC
901
ACACAAGAGTGTCATACCCCGGTCCGTCTCCATACACAATGGTCTTCAAAATGACCTTCG
Hind II
AAGTGAACCACAGTCAACCCTAGCTTACGGGAGCCCTGACATTTCATCTCTTTTGAGAGA
961
TTCACTTGGTGTCAGTTGGGATCGAATGCCCTCGGGACTGTAAAGTAGAGAAAACTCTCT
PRIMER 3 5'CTCAAGAAAATACAGACGAA3' T_m=54°C
Pst I
TTCATCTTCAACTCAAGAAAATACAGACGAAGAATCTGGCCCGTGCTGCAGCAAGACTCT
1021
AAGTAGAAGTTGAGTTCTTTTATGTCTGCTTCTTAGACCGGGCACGACGTCGTTCTGAGA
CAGCCCAGGTGTGCCACAGCCTCAATCTGAATATGACCCCAGCCCGACCTCTCCACCTGA
1081
GTCGGGTCCACACGGTGTCGGAGTTAGACTTATACTGGGGTCGGGCTGGAGAGGTGGACT
TTCGGACACTGAGTCATGTGATAGCCAGGTTCGTCCAGAAAGTACTGACTCCGACACCCA
1141
AAGCCTGTGACTCAGTACACTATCGGTCCAAGCAGGTCTTTCATGACTGAGGCTGTGGGT
TGCAGAAGACGATGATGTTCCTGAAGCACCCCAAGCAGCAAGCCAAACTCAGCCAACAAC
1201
ACGTCTTCTGCTACTACAAGGACTTCGTGGGGTTCGTCGTTCGGTTTGAGTCGGTTGTTG
Pst I
TACCCAAGAGACACAGAGTTGCTGCAGTGTCTACAATCCAGCTACCACTCAGACGGGATT
1261
ATGGGTTCTCTGTGTCTCAACGACGTCACAGATGTTAGGTCGATGGTGAGTCTGCCCTAA
CTACTATAACCAGCAGTCAAGTGATAATTTTGCCTCTCGTATAGATGCATCTGCCACATC
1321
GATGATATTGGTCGTCAGTTCACTATTAAAACGGAGAGCATATCTACGTAGACGGTGTAG
PRIMER 14 T_m=58°C 3'TATCTACGTAGACGGTGTAG5'
AACCAGTTATGGGTATCCAGGCCCAGTCACTAATCACGGCTTTTCAACTTCTGTCTTTTC
1381
TTGGTCAATACCCATAGGTCCGGGTCAGTGATTAGTGCCGAAAAGTTGAAGACAGAAAAG


Pvu II
CCATACAGTACCCACAACGGCCACTGGCAGCAGCAGCTGCATCAAAATATGTATGGGGGT
1441 ---------+---------+---------+---------+---------+---------+ 1500
GGTATGTCATGGGTGTTGCCGGTGACCGTCGTCGTCGACGTAGTTTTATACATACCCCCA
GGTCAGCAGACGACCACATACGGCAGCTATGTGGGGGGTTACTCAGATGCTAATGGCCAG
1501 ---------+---------+---------+---------+---------+---------+ 1560
CCAGTCGTCTGCTGGTGTATGCCGTCGATACACCCCCCAATGAGTCTACGATTACCGGTC
TCTGTTGGTGCAAGTACCAGTTACTACACCAGTAAGCCTGCGACCAGCCGCAGCAGCACC
1561 ---------+---------+---------+---------+---------+---------+ 1620
AGACAACCACGTTCATGGTCAATGATGTGGTCATTCGGACGCTGGTCGGCGTCGTCGTGG
Pst I
GCTCTGCAGTGGACCACTTTGTTTCCACCGGCCAGCTCATCTGCTTCCAGCAGCCAAGTT
1621 ---------+---------+---------+---------+---------+---------+ 1680
CGAGACGTCACCTGGTGAAACAAAGGTGGCCGGTCGAGTAGACGAAGGTCGTCGGTTCAA
PRIMER 8 3'TTCAA
TCTGATTATTTCAGCCCCTTCCCTTATTTTCCTGGCAGTTCTACAGTTCCTCAGGCCTTC
1681 ---------+---------+---------+---------+---------+---------+ 1740
AGACTAATAAAGTCGGGGAAGGGAATAAAAGGACCGTCAAGATGTCAAGGAGTCCGGAAG
AGACTAATAAAGTCG5' T_m=52°C
GAGCCTTTAGCGCCATCTACCCCTAGCTTGCTGGATGAACTGCTTGATAGAGACAGTGGA
1741 ---------+---------+---------+---------+---------+---------+ 1800
CTCGGAAATCGCGGTAGATGGGGATCGAACGACCTACTTGACGAACTATCTCTGTCACCT
CTTGTAAGCCAGCAGCAGGCACCTGCACCTCCCCAAAATGACCAGGGGGGGCCTCCTCAG
1801 ---------+---------+---------+---------+---------+---------+ 1860
GAACATTCGGTCGTCGTCCGTGGACGTGGAGGGGTTTTACTGGTCCCCCCCGGAGGAGTC
TATGTGCCCGTTGCTCAAGAGCAACAGCAGTCCAGTACAGACCCGCTGTCTGATGAGATG
1861 ---------+---------+---------+---------+---------+---------+ 1920
ATACACGGGCAACGAGTTCTCGTTGTCGTCAGGTCATGTCTGGGCGACAGACTACTCTAC
AGGAGAATTTTTTGAGTTTTTGACAGTGTGAACCCCGTTTGAACCAGACCATAGCCATAT
1921 ---------+---------+---------+---------+---------+---------+ 1980
TCCTCTTAAAAAACTCAAAAACTGTCACACTTGGGGCAAACTTGGTCTGGTATCGGTATA
TTTGGGAGTGTAAACAGAATCCTCTGTAAACTGCCTTTTTTATGTGAAATGCGTCTCTCG
1981 ---------+---------+---------+---------+---------+---------+ 2040
AAACCCTCACATTTGTCTTAGGAGACATTTGACGGAAAAAATACACTTTACGCAGAGAGC
CCAAAACTATCTGCTCTTGTTTTGACAAAAAGTGGAAGTGTGTGTTTATATGATTCACAA
2041 ---------+---------+---------+---------+---------+---------+ 2100
GGTTTTGATAGACGAGAACAAAACTGTTTTTCACCTTCACACACAAATATACTAAGTGTT
AACTGTAGTTTAACAGTAAAAGTTAATTGTGGTTTGTGGTATGACACACAGTTCTGGAAA
2101 ---------+---------+---------+---------+---------+---------+ 2160
TTGACATCAAATTGTCATTTTCAATTAACACCAAACACCATACTGTGTGTCAAGACCTTT
CTGACCAAACAGTGTGTTCTTATGTGCACTTATTTTGTATTTCCCTTATGCCTGCCAGAG
2161 ---------+---------+---------+---------+---------+---------+ 2220
GACTGGTTTGTCACACAAGAATACACGTGAATAAAACATAAAGGGAATACGGACGGTCTC


PRIMER 1 5'TCAACACTATAA
TGCTCAATAAAAGTTACACTCAAGTGCCACTGAGAGTTTTTTCTTTTTTCAACACTATAA
2221 ---------+---------+---------+---------+---------+---------+ 2280
ACGAGTTATTTTCAATGTGAGTTCACGGTGACTCTCAAAAAAGAAAAAAGTTGTGATATT
AAGCATCT3' T_m=52°C
Pvu II
AAGCATCTTTGGGGACTCATGCATAAGCACTCTGCTACAGCTGACATGCTACCCATTTCA
2281 ---------+---------+---------+---------+---------+---------+ 2340
TTCGTAGAAACCCCTGAGTACGTATTCGTGAGACGATGTCGACTGTACGATGGGTAAAGT
GCATTTAGCCCTCTGAAATTCTTGGGTGGAGTGGTGGCTACACCCCAGGGGCCAGTATTG
2341 ---------+---------+---------+---------+---------+---------+ 2400
CGTAAATCGGGAGACTTTAAGAACCCACCTCACCACCGATGTGGGGTCCCCGGTCATAAC
TGGGATGCCACTACTACTCCTGTCCCTTTCCAGGCTTTCCCCTGCCCTCCGGCAATAACT
2401 ---------+---------+---------+---------+---------+---------+ 2460
ACCCTACGGTGATGATGAGGACAGGGAAAGGTCCGAAAGGGGACGGGAGGCCGTTATTGA
CCTATAAATCAGTTTGTTAATTATCAGCCTGCGCCTGCGATGGATAAGATATGGAGACCT
2461 ---------+---------+---------+---------+---------+---------+ 2520
GGATATTTAGTCAAACAATTAATAGTCGGACGCGGACGCTACCTATTCTATACCTCTGGA
Kpn I
GCCTTTGAGGACTACCGTGGGGTAACTATGAGACCGGCCTTACCGGAGGTACCTGAAACA
2521 ---------+---------+---------+---------+---------+---------+ 2580
CGGAAACTCCTGATGGCACCCCATTGATACTCTGGCCGGAATGGCCTCCATGGACTTTGT
AGTTTGAGACCAGCTTTACAGGAGCTACCCGAGCCATCATCTCCTCAAAGTCAGAGTTCT
2581 ---------+---------+---------+---------+---------+---------+ 2640
TCAAACTCTGGTCGAAATGTCCTCGATGGGCTCGGTAGTAGAGGAGTTTCAGTCTCAAGA
GTGGATGATGACACAGATTCTAAAGAAGATGTTACAGAGACCTTAGAATGTGCTCAGGCT
2641 ---------+---------+---------+---------+---------+---------+ 2700
CACCTACTACTGTGTCTAAGATTTCTTCTACAATGTCTCTGGAATCTTACACGAGTCCGA
CTGACAGATCTGAAGTGGGGTACAGTGGACCCTTCCGAGAAGACATCTCAGTTACTGCTT
2701 ---------+---------+---------+---------+---------+---------+ 2760
GACTGTCTAGACTTCACCCCATGTCACCTGGGAAGGCTCTTCTGTAGAGTCAATGACGAA
GCACTTCAAGCGGCGAGACCAGCATTTGCAGAAGGCAAAGGAACAGCTATACGCGCCTAC
2761 ---------+---------+---------+---------+---------+---------+ 2820
CGTGAAGTTCGCCGCTCTGGTCGTAAACGTCTTCCGTTTCCTTGTCGATATGCGCGGATG
AGACTTATTGAAGCTAAAGACTTGGGCAAGCTAGTAGTGAGAGGACCCAAAAAAGACAAG
2821 ---------+---------+---------+---------+---------+---------+ 2880
TCTGAATAACTTCGATTTCTGAACCCGTTCGATCATCACTCTCCTGGGTTTTTTCTGTTC
Kpn I
CGAGATCCAGACTTCTACATCAAAAAGGTACCTTTACATTTTTCTTAAAATTAAAAGTAA
2881 ---------+---------+---------+---------+---------+---------+ 2940
GCTCTAGGTCTGAAGATGTAGTTTTTCCATGGAAATGTAAAAAGAATTTTAATTTTCATT

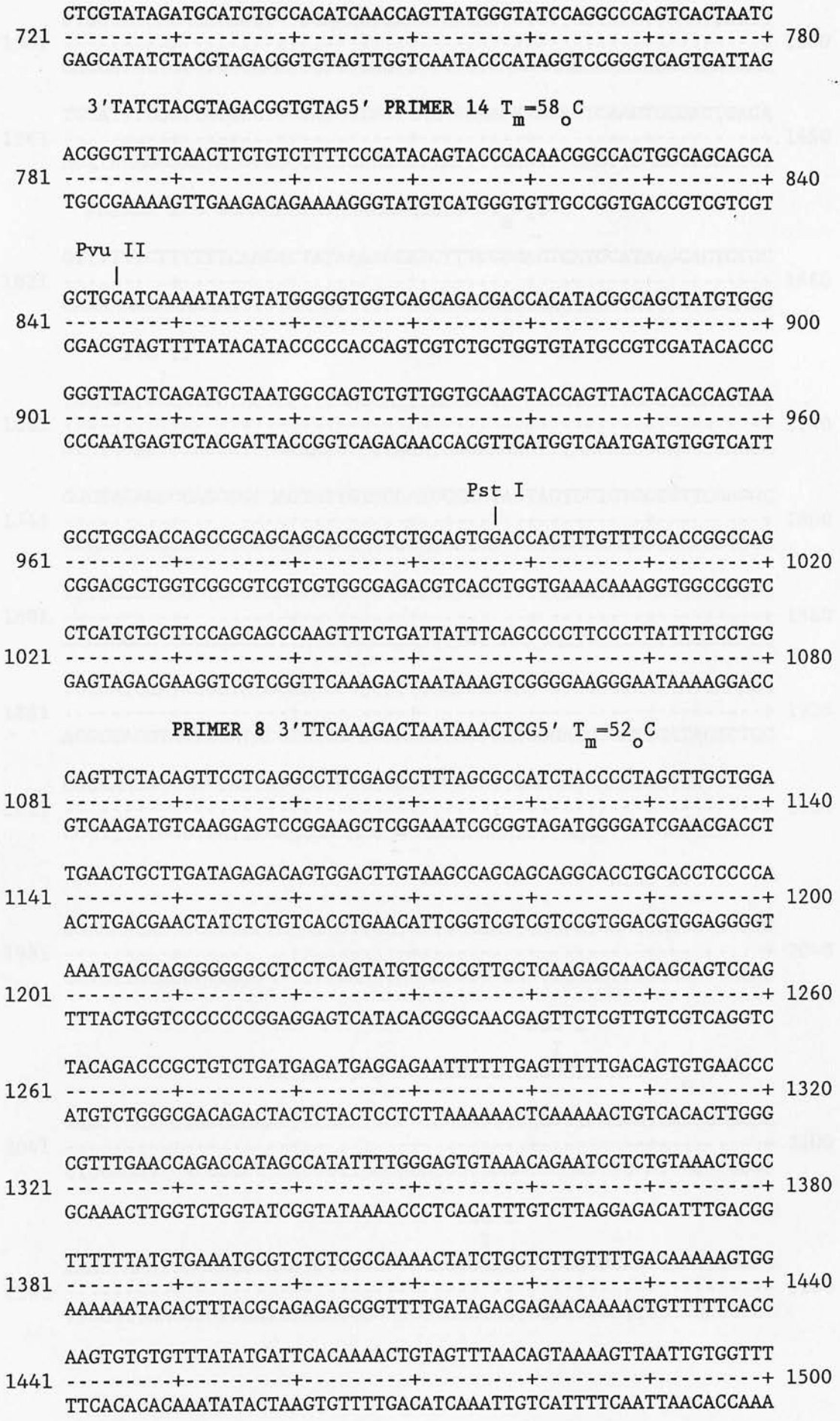
```
     CTTTTTAGGTTGCATGAAGCTAACACTTTTTTTCTCTTTTTATAGATACAGCAGAGTACC
2941 ---------+---------+---------+---------+---------+---------+ 3000
     GAAAAATCCAACGTACTTCGATTGTGAAAAAAAGAGAAAAATATCTATGTCGTCTCATGG

                                        Sph I
                                          |
     ACAAAGGCAGAGGTGCGCAGATGGAGTACAAGATTGCATGCTAAACTCTTTAGAGAAAAT
3001 ---------+---------+---------+---------+---------+---------+ 3060
     TGTTTCCGTCTCCACGCGTCTACCTCATGTTCTAACGTACGATTTGAGAAATCTCTTTTA

          PRIMER 2 5'AGATGATGAAGGTAAAGGAC3' T_m=56°C

     TACCACAAACGAATACTTCAGATGATGAAGGTAAAGGACTTCCCCGTATAAATATACCCC
3061 ---------+---------+---------+---------+---------+---------+ 3120
     ATGGTGTTTGCTTATGAAGTCTACTACTTCCATTTCCTGAAGGGGCATATTTATATGGGG

     CCGTGCCCTTTGGGAAAAGAGTACCTAAACACTTAATACTCTTTTTTACACTCCAGGAGC
3121 ---------+---------+---------+---------+---------+---------+ 3180
     GGCACGGGAAACCCTTTTCTCATGGATTTGTGAATTATGAGAAAAAATGTGAGGTCCTCG

     TGGAAGATAAAACAAAAGAACTGAAATTTTTACAGGAGAGTACCTTGCGGCTTTTAAAAC
3181 ---------+---------+---------+---------+---------+---------+ 3240
     ACCTTCTATTTTGTTTTCTTGACTTTAAAAATGTCCTCTCATGGAACGCCGAAAATTTTG

     AGAATGAGAAACTAAGGGCACGGCTGGCTTCCCTGCATGTAAAAACATCAGCTACGAACT
3241 ---------+---------+---------+---------+---------+---------+ 3300
     TCTTACTCTTTGATTCCCGTGCCGACCGAAGGGACGTACATTTTTGTAGTCGATGCTTGA

     AAACCCCCGCCCCATTCTACCCTCAATCTATAAATAAAGTTTTGTTTGCACCCATTTTTA
3301 ---------+---------+---------+---------+---------+---------+ 3360
     TTTGGGGGCGGGGTAAGATGGGAGTTAGATATTTATTTCAAAACAAACGTGGGTAAAAAT

     TGTGTTTGTGTGTATGTTTAGGGGTGGGTATTGGTTTtAAAAAGTAAAAGATTAGTATTG
3361 ---------+---------+---------+---------+---------+---------+ 3420
     ACACAAACACACATACAAATCCCCACCCATAACCAAAATTTTTCATTTTCTAATCATAAC

     TGACATCTGGGACAAAGAAACTATTTTTAACCTTTTCAGGTATAGACTTTACCCTATTGA
3421 ---------+---------+---------+---------+---------+---------+ 3480
     ACTGTAGACCCTGTTTCTTTGATAAAAATTGGAAAAGTCCATATCTGAAATGGGATAACT

     CACACTCTTAAGTAACTATTTAAGATATTAATGTTTTGTAAGCCTTGCCTACGTACCGGG
3481 ---------+---------+---------+---------+---------+---------+ 3540
     GTGTGAGAATTCATTGATAAATTCTATAATTACAAAACATTCGGAACGGATGCATGGCCC

     AGTGTAAATAACTTTAAGAGCTTCCCCCGTAACTCGTATGTATTCTGTTTATTTAGTGTG
3541 ---------+---------+---------+---------+---------+---------+ 3600
     TCACATTTATTGAAATTCTCGAAGGGGGCATTGAGCATACATAAGACAAATAAATCACAC

     ACTGTATTTTAAAAATATTTAACTGTGGCTGGGCTCTGAGATGTTGGCAGAAATGTTGTG
3601 ---------+---------+---------+---------+---------+---------+ 3660
     TGACATAAAATTTTTATAAATTGACACCGACCCGAGACTCTACAACCGTCTTTACAACAC

     Pst I  PRIMER 4 3'GACACCGACCCGAGACTCTA5' T_m=54°C
       |
     GCCTGCAGTGAACATGCTGCTTCCCCGTAAAGCACTTTTGGTAGACATTTTTTTTATTTT
3661 ---------+---------+---------+---------+---------+---------+ 3720
     CGGACGTCACTTGTACGACGAAGGGGCATTTCGTGAAAACCATCTGTAAAAAAAATAAAA
```

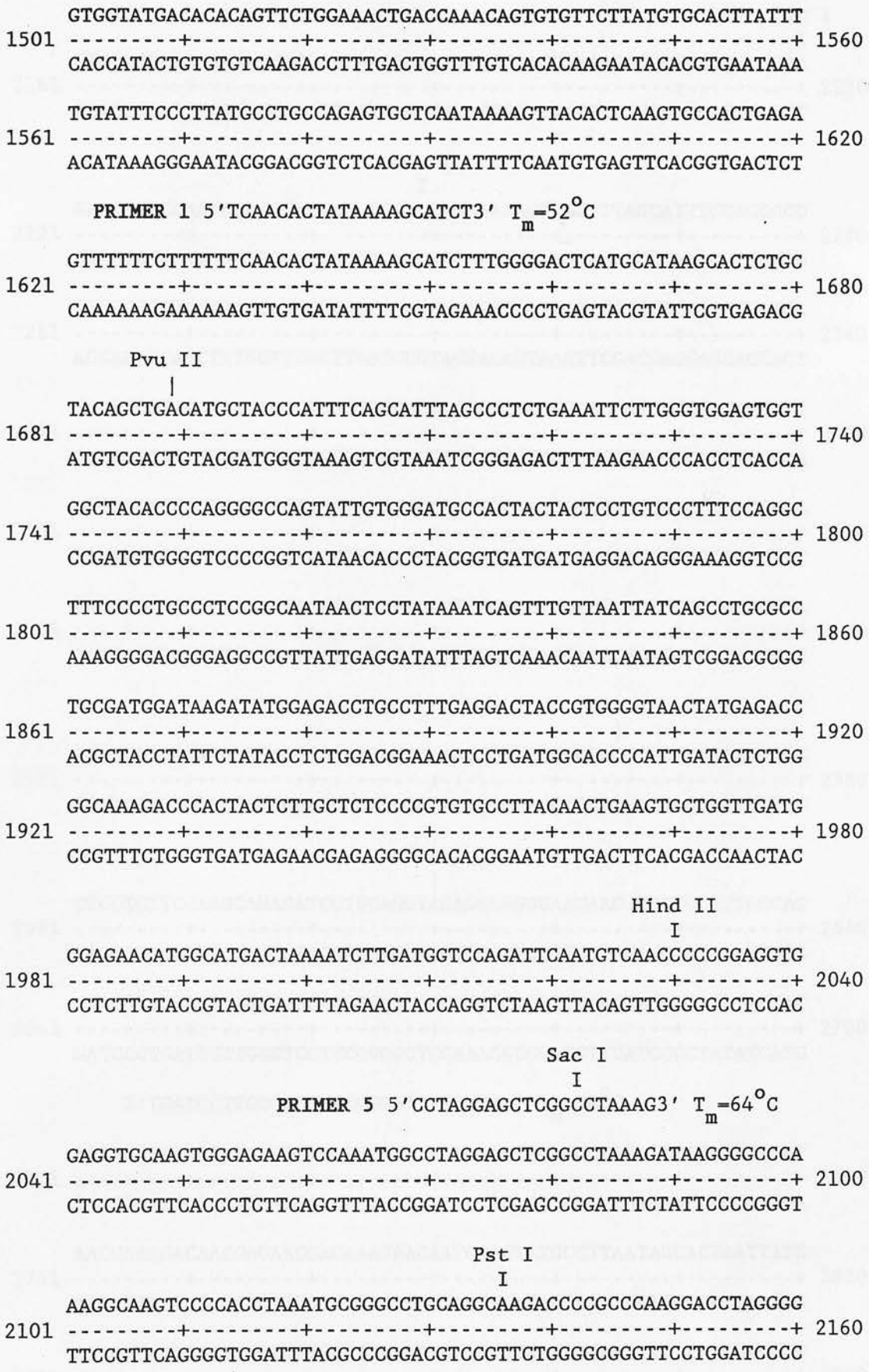
```
    GGCAGCCACAAATTTAATGATAGCTGCCTTTGCCCTGGGCTGCTTGGCTTTCTATAATGC
3721 ---------+---------+---------+---------+---------+---------+ 3780
    CCGTCGGTGTTTAAATTACTATCGACGGAAACGGGACCCGACGAACCGAAAGATATTACG

    CGAGATCTCGGACCTCTCTGCGGACGCCACGCCCCTTTCCCCGGGTCCGCCC
3781 ---------+---------+---------+---------+---------+-- 3832
    GCTCTAGAGCCTGGAGAGACGCCTGCGGTGCGGGGAAAGGGGCCCAGGCGGG
```

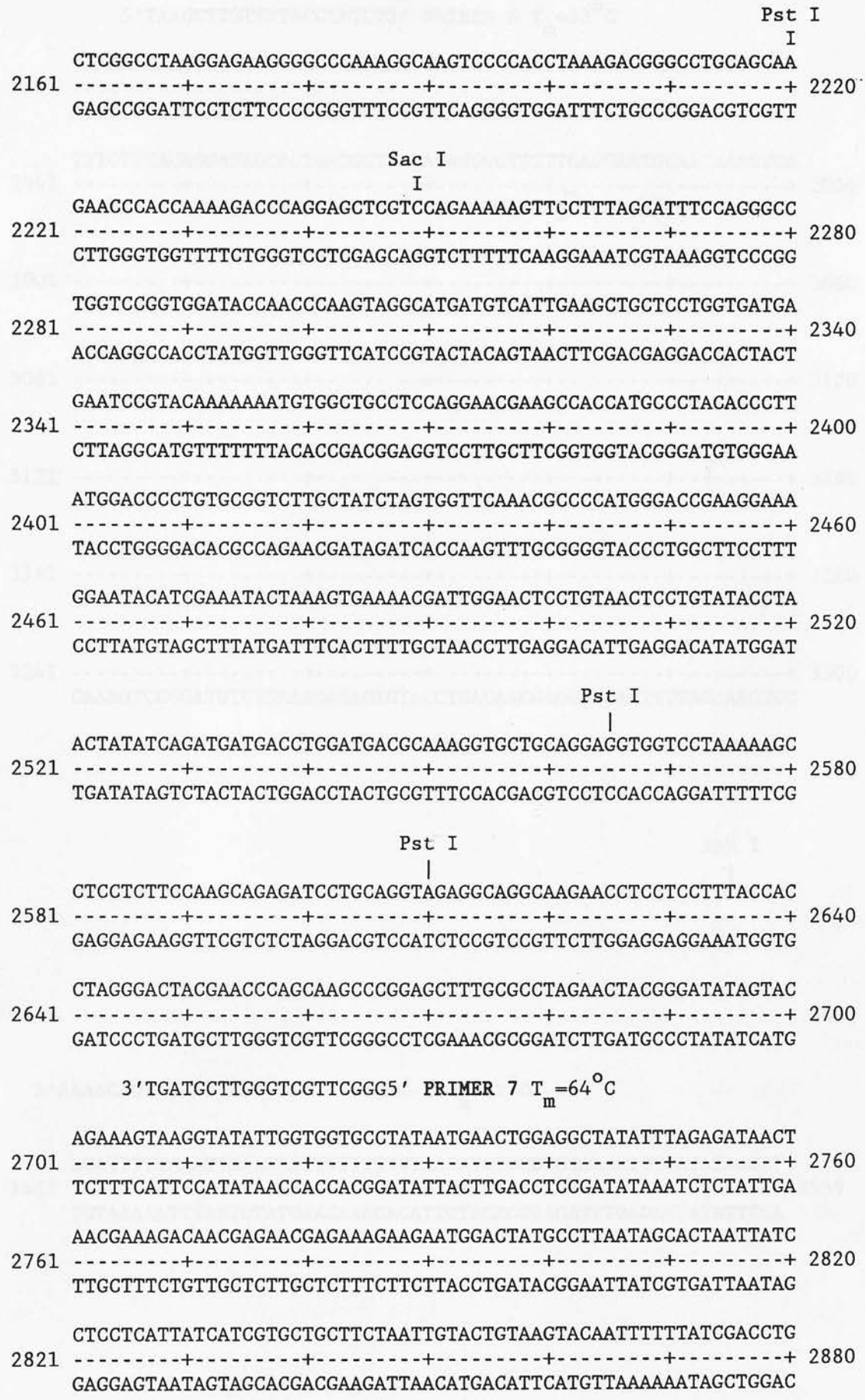
PRIMER 15 3'CTGCGGTGCGGGGAAAGGGG5' T_m=70°C

Appendix 3.2 The complete sequence of VIR-1

```
    AGCTTTGAACCAAGGGAGAATGTTTAGCGGAGGGAGACAAAACATGATGAATGCCGCCAA
  1 ---------+---------+---------+---------+---------+---------+  6
    TCGAAACTTGGTTCCCTCTTACAAATCGCCTCCCTCTGTTTTGTACTACTTACGGCGGTT

    GAAGGTTTTGACAGTGTACTCTAGCCTCAGAGACGATGGAGAAATCAGTCCAGAAGTAAA
 61 ---------+---------+---------+---------+---------+---------+ 120
    CTTCCAAAACTGTCACATGAGATCGGAGTCTCTGCTACCTCTTTAGTCAGGTCTTCATTT

    GGCTTACATGGCTTACATTTTCCAGTGCCTGACATTGAGCAGGTATTCCAACCACTGTTT
121 ---------+---------+---------+---------+---------+---------+ 180
    CCGAATGTACCGAATGTAAAAGGTCACGGACTGTAACTCGTCCATAAGGTTGGTGACAAA

    AAGTTGGAGCAAGAAATCAGAAAAGGTAAAGCCACCTTGACACAGTCTCTGCTGTTTGCA
181 ---------+---------+---------+---------+---------+---------+ 240
    TTCAACCTCGTTCTTTAGTCTTTTCCATTTCGGTGGAACTGTGTCAGAGACGACAAACGT

    CCAAGGAAAGATGTCCCAATAAGACTGTGTTCTCACAGTATGGGGCCAGGCAGAGGTATG
241 ---------+---------+---------+---------+---------+---------+ 300
    GGTTCCTTTCTACAGGGTTATTCTGACACAAGAGTGTCATACCCCGGTCCGTCTCCATAC

                                                Hind II
                                                   |
    TGTTACCAGAAGTTTTACTGGAAGCAAGTGAACCACAGTCAACCCTAGCTTACGGGAGCC
301 ---------+---------+---------+---------+---------+---------+ 360
    ACAATGGTCTTCAAAATGACCTTCGTTCACTTGGTGTCAGTTGGGATCGAATGCCCTCGG

                    PRIMER 3 T_m=54°C 5'CTCAAGAAAATACAGACGAA3'

    CTGACATTTCATCTCTTTTGAGAGATTCATCTTCAACTCAAGAAAATACAGACGAAGAAT
361 ---------+---------+---------+---------+---------+---------+ 420
    GACTGTAAAGTAGAGAAAACTCTCTAAGTAGAAGTTGAGTTCTTTTATGTCTGCTTCTTA

                  PstI
                   |
    CTGGCCCGTGCTGCAGCAAGACTCTCAGCCCAGGTGTGCCACAGCCTCAATCTGAATATG
421 ---------+---------+---------+---------+---------+---------+ 480
    GACCGGGCACGACGTCGTTCTGAGAGTCGGGTCCACACGGTGTCGGAGTTAGACTTATAC

    ACCCCAGCCCGACCTCTCCACCTGATTCGGACACTGAGTCATGTGATAGCCAGGTTCGTC
481 ---------+---------+---------+---------+---------+---------+ 540
    TGGGGTCGGGCTGGAGAGGTGGACTAAGCCTGTGACTCAGTACACTATCGGTCCAAGCAG

    CAGAAAGTACTGACTCCGACACCCATGCAGAAGACGATGATGTTCCTGAAGCACCCCAAG
541 ---------+---------+---------+---------+---------+---------+ 600
    GTCTTTCATGACTGAGGCTGTGGGTACGTCTTCTGCTACTACAAGGACTTCGTGGGGTTC

                                                      Pst I
                                                       |
    CAGCAAGCCAAACTCAGCCAACAACTACCCAAGAGACACAGAGTTGCTGCAGTGTCTACA
601 ---------+---------+---------+---------+---------+---------+ 660
    GTCGTTCGGTTTGAGTCGGTTGTTGATGGGTTCTCTGTGTCTCAACGACGTCACAGATGT

    ATCCAGCTACCACTCAGACGGGATTCTACTATAACCAGCAGTCAAGTGATAATTTTGCCT
661 ---------+---------+---------+---------+---------+---------+ 720
    TAGGTCGATGGTGAGTCTGCCCTAAGATGATATTGGTCGTCAGTTCACTATTAAAACGGA
```


CTCGTATAGATGCATCTGCCACATCAACCAGTTATGGGTATCCAGGCCCAGTCACTAATC
721
GAGCATATCTACGTAGACGGTGTAGTTGGTCAATACCCATAGGTCCGGGTCAGTGATTAG
780
3'TATCTACGTAGACGGTGTAG5' PRIMER 14 T_m=58 ^{o}C
ACGGCTTTTCAACTTCTGTCTTTTCCCATACAGTACCCACAACGGCCACTGGCAGCAGCA
781
TGCCGAAAAGTTGAAGACAGAAAAGGGTATGTCATGGGTGTTGCCGGTGACCGTCGTCGT
840
Pvu II
GCTGCATCAAAATATGTATGGGGGTGGTCAGCAGACGACCACATACGGCAGCTATGTGGG
841
CGACGTAGTTTTATACATACCCCCACCAGTCGTCTGCTGGTGTATGCCGTCGATACACCC
900
GGGTTACTCAGATGCTAATGGCCAGTCTGTTGGTGCAAGTACCAGTTACTACACCAGTAA
901
CCCAATGAGTCTACGATTACCGGTCAGACAACCACGTTCATGGTCAATGATGTGGTCATT
960
Pst I
GCCTGCGACCAGCCGCAGCAGCACCGCTCTGCAGTGGACCACTTTGTTTCCACCGGCCAG
961
CGGACGCTGGTCGGCGTCGTCGTGGCGAGACGTCACCTGGTGAAACAAAGGTGGCCGGTC
1020
CTCATCTGCTTCCAGCAGCCAAGTTTCTGATTATTTCAGCCCCTTCCCTTATTTTCCTGG
1021
GAGTAGACGAAGGTCGTCGGTTCAAAGACTAATAAAGTCGGGGAAGGGAATAAAAGGACC
1080
PRIMER 8 3'TTCAAAGACTAATAAACTCG5' T_m=52 ^{o}C
CAGTTCTACAGTTCCTCAGGCCTTCGAGCCTTTAGCGCCATCTACCCCTAGCTTGCTGGA
1081
GTCAAGATGTCAAGGAGTCCGGAAGCTCGGAAATCGCGGTAGATGGGGATCGAACGACCT
1140
TGAACTGCTTGATAGAGACAGTGGACTTGTAAGCCAGCAGCAGGCACCTGCACCTCCCCA
1141
ACTTGACGAACTATCTCTGTCACCTGAACATTCGGTCGTCGTCCGTGGACGTGGAGGGGT
1200
AAATGACCAGGGGGGGCCTCCTCAGTATGTGCCCGTTGCTCAAGAGCAACAGCAGTCCAG
1201
TTTACTGGTCCCCCCCGGAGGAGTCATACACGGGCAACGAGTTCTCGTTGTCGTCAGGTC
1260
TACAGACCCGCTGTCTGATGAGATGAGGAGAATTTTTTGAGTTTTTGACAGTGTGAACCC
1261
ATGTCTGGGCGACAGACTACTCTACTCCTCTTAAAAAACTCAAAAACTGTCACACTTGGG
1320
CGTTTGAACCAGACCATAGCCATATTTTGGGAGTGTAAACAGAATCCTCTGTAAACTGCC
1321
GCAAACTTGGTCTGGTATCGGTATAAAACCCTCACATTTGTCTTAGGAGACATTTGACGG
1380
TTTTTTATGTGAAATGCGTCTCTCGCCAAAACTATCTGCTCTTGTTTTGACAAAAAGTGG
1381
AAAAAATACACTTTACGCAGAGAGCGGTTTTGATAGACGAGAACAAAACTGTTTTTCACC
1440
AAGTGTGTGTTTATATGATTCACAAAACTGTAGTTTAACAGTAAAAGTTAATTGTGGTTT
1441
TTCACACACAAATATACTAAGTGTTTTGACATCAAATTGTCATTTTCAATTAACACCAAA
1500


GTGGTATGACACACAGTTCTGGAAACTGACCAAACAGTGTGTTCTTATGTGCACTTATTT
1501 + 1560
CACCATACTGTGTGTCAAGACCTTTGACTGGTTTGTCACACAAGAATACACGTGAATAAA
TGTATTTCCCTTATGCCTGCCAGAGTGCTCAATAAAAGTTACACTCAAGTGCCACTGAGA
1561 + 1620
ACATAAAGGGAATACGGACGGTCTCACGAGTTATTTTCAATGTGAGTTCACGGTGACTCT
PRIMER 1 5'TCAACACTATAAAAGCATCT3' T_m=52°C
GTTTTTTCTTTTTTCAACACTATAAAAGCATCTTTGGGGACTCATGCATAAGCACTCTGC
1621 + 1680
CAAAAAAGAAAAAAGTTGTGATATTTTCGTAGAAACCCCTGAGTACGTATTCGTGAGACG
Pvu II
TACAGCTGACATGCTACCCATTTCAGCATTTAGCCCTCTGAAATTCTTGGGTGGAGTGGT
1681 + 1740
ATGTCGACTGTACGATGGGTAAAGTCGTAAATCGGGAGACTTTAAGAACCCACCTCACCA
GGCTACACCCCAGGGGCCAGTATTGTGGGATGCCACTACTACTCCTGTCCCTTTCCAGGC
1741 + 1800
CCGATGTGGGGTCCCCGGTCATAACACCCTACGGTGATGATGAGGACAGGGAAAGGTCCG
TTTCCCCTGCCCTCCGGCAATAACTCCTATAAATCAGTTTGTTAATTATCAGCCTGCGCC
1801 + 1860
AAAGGGGACGGGAGGCCGTTATTGAGGATATTTAGTCAAACAATTAATAGTCGGACGCGG
TGCGATGGATAAGATATGGAGACCTGCCTTTGAGGACTACCGTGGGGTAACTATGAGACC
1861 + 1920
ACGCTACCTATTCTATACCTCTGGACGGAAACTCCTGATGGCACCCCATTGATACTCTGG
GGCAAAGACCCACTACTCTTGCTCTCCCCGTGTGCCTTACAACTGAAGTGCTGGTTGATG
1921 + 1980
CCGTTTCTGGGTGATGAGAACGAGAGGGGCACACGGAATGTTGACTTCACGACCAACTAC
Hind II
GGAGAACATGGCATGACTAAAATCTTGATGGTCCAGATTCAATGTCAACCCCCGGAGGTG
1981 + 2040
CCTCTTGTACCGTACTGATTTTAGAACTACCAGGTCTAAGTTACAGTTGGGGGCCTCCAC
Sac I
PRIMER 5 5'CCTAGGAGCTCGGCCTAAAG3' T_m=64°C
GAGGTGCAAGTGGGAGAAGTCCAAATGGCCTAGGAGCTCGGCCTAAAGATAAGGGGCCCA
2041 + 2100
CTCCACGTTCACCCTCTTCAGGTTTACCGGATCCTCGAGCCGGATTTCTATTCCCCGGGT
Pst I
AAGGCAAGTCCCCACCTAAATGCGGGCCTGCAGGCAAGACCCCGCCCAAGGACCTAGGGG
2101 + 2160
TTCCGTTCAGGGGTGGATTTACGCCCGGACGTCCGTTCTGGGGCGGGTTCCTGGATCCCC


Pst I
CTCGGCCTAAGGAGAAGGGGCCCAAAGGCAAGTCCCCACCTAAAGACGGGCCTGCAGCAA
2161
2220
GAGCCGGATTCCTCTTCCCCGGGTTTCCGTTCAGGGGTGGATTTCTGCCCGGACGTCGTT
Sac I
GAACCCACCAAAAGACCCAGGAGCTCGTCCAGAAAAAGTTCCTTTAGCATTTCCAGGGCC
2221
2280
CTTGGGTGGTTTTCTGGGTCCTCGAGCAGGTCTTTTTCAAGGAAATCGTAAAGGTCCCGG
TGGTCCGGTGGATACCAACCCAAGTAGGCATGATGTCATTGAAGCTGCTCCTGGTGATGA
2281
2340
ACCAGGCCACCTATGGTTGGGTTCATCCGTACTACAGTAACTTCGACGAGGACCACTACT
GAATCCGTACAAAAAAATGTGGCTGCCTCCAGGAACGAAGCCACCATGCCCTACACCCTT
2341
2400
CTTAGGCATGTTTTTTTACACCGACGGAGGTCCTTGCTTCGGTGGTACGGGATGTGGGAA
ATGGACCCCTGTGCGGTCTTGCTATCTAGTGGTTCAAACGCCCCATGGGACCGAAGGAAA
2401
2460
TACCTGGGGACACGCCAGAACGATAGATCACCAAGTTTGCGGGGTACCCTGGCTTCCTTT
GGAATACATCGAAATACTAAAGTGAAAACGATTGGAACTCCTGTAACTCCTGTATACCTA
2461
2520
CCTTATGTAGCTTTATGATTTCACTTTTGCTAACCTTGAGGACATTGAGGACATATGGAT
Pst I
ACTATATCAGATGATGACCTGGATGACGCAAAGGTGCTGCAGGAGGTGGTCCTAAAAAGC
2521
2580
TGATATAGTCTACTACTGGACCTACTGCGTTTCCACGACGTCCTCCACCAGGATTTTTCG
Pst I
CTCCTCTTCCAAGCAGAGATCCTGCAGGTAGAGGCAGGCAAGAACCTCCTCCTTTACCAC
2581
2640
GAGGAGAAGGTTCGTCTCTAGGACGTCCATCTCCGTCCGTTCTTGGAGGAGGAAATGGTG
CTAGGGACTACGAACCCAGCAAGCCCGGAGCTTTGCGCCTAGAACTACGGGATATAGTAC
2641
2700
GATCCCTGATGCTTGGGTCGTTCGGGCCTCGAAACGCGGATCTTGATGCCCTATATCATG
3'TGATGCTTGGGTCGTTCGGG5' PRIMER 7 T_m=64°C
AGAAAGTAAGGTATATTGGTGGTGCCTATAATGAACTGGAGGCTATATTTAGAGATAACT
2701
2760
TCTTTCATTCCATATAACCACCACGGATATTACTTGACCTCCGATATAAATCTCTATTGA
AACGAAAGACAACGAGAACGAGAAAGAAGAATGGACTATGCCTTAATAGCACTAATTATC
2761
2820
TTGCTTTCTGTTGCTCTTGCTCTTTCTTCTTACCTGATACGGAATTATCGTGATTAATAG
CTCCTCATTATCATCGTGCTGCTTCTAATTGTACTGTAAGTACAATTTTTTATCGACCTG
2821
2880
GAGGAGTAATAGTAGCACGACGAAGATTAACATGACATTCATGTTAAAAAATAGCTGGAC

5'TAACCTTGTTTTACCTATGT3' **PRIMER 6** T_m=52°C

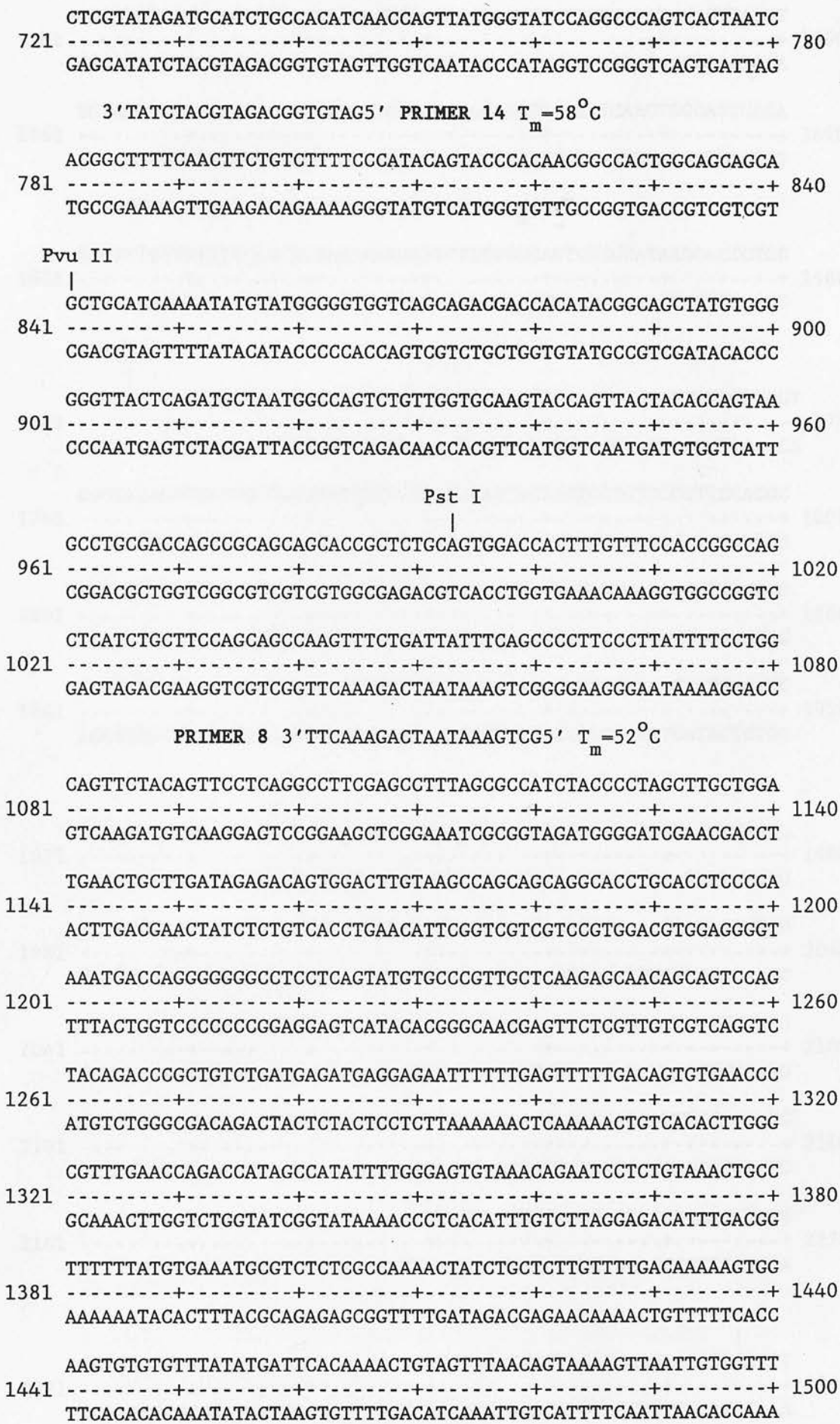
```
     TAACAATAACCTTGTTTTACCTATGTAACCTTAATTGTATTTTATATTAACATCTCATTA
2881 ---------+---------+---------+---------+---------+---------+ 2940
     ATTGTTATTGGAACAAAATGGATACATTGGAATTAACATAAAATATAATTGTAGAGTAAT

     TTTCTTTACAGGATAGCACTGATGGTTAAACAGGGCTTCTTCAGCAATGCAACAAACTCA
2941 ---------+---------+---------+---------+---------+---------+ 3000
     AAAGAAATGTCCTATCGTGACTACCAATTTGTCCCGAAGAAGTCGTTACGTTGTTTGAGT

     ACAGCTTCCCATTAATGGACATGACCAGTTATCGCGCCTAGACTCCTAAATACAGGTGAT
3001 ---------+---------+---------+---------+---------+---------+ 3060
     TGTCGAAGGGTAATTACCTGTACTGGTCAATAGCGCGGATCTGAGGATTTATGTCCACTA

     GTTGGCACAAGTTTAGGCCTACTGCCTGATTTTGCTAGACACTTTGCCTCAATTTTACTT
3061 ---------+---------+---------+---------+---------+---------+ 3120
     CAACCGTGTTCAAATCCGGATGACGGACTAAAACGATCTGTGAAACGGAGTTAAAATGAA

     TTGCTGCACTTTAAACTGTTTCAGGGCTAGAGACTTTGCTGAGATGGACTGTTGCTGCAC
3121 ---------+---------+---------+---------+---------+---------+ 3180
     AACGACGTGAAATTTGACAAAGTCCCGATCTCTGAAACGACTCTACCTGACAACGACGTG

     TTTAAACTGTTTCAGGGCTAGAGACTTTGTCTGAGATGGACTGTTGCTGCACTTTAAACT
3181 ---------+---------+---------+---------+---------+---------+ 3240
     AAATTTGACAAAGTCCCGATCTCTGAAACAGACTCTACCTGACAACGACGTGAAATTTGA

     GTTTCAGGGCTAGAGACTTTGTCTGAGATGGACTGTTGCTGCACTTTAAACTGTTTCAGG
3241 ---------+---------+---------+---------+---------+---------+ 3300
     CAAAGTCCCGATCTCTGAAACAGACTCTACCTGACAACGACGTGAAATTTGACAAAGTCC

     GCTAGAGACTTTGTCTGAGATGGACTGTTGCTGCACTTTAAACTTTTTCAGGGCTAGAGA
3301 ---------+---------+---------+---------+---------+---------+ 3360
     CGATCTCTGAAACAGACTCTACCTGACAACGACGTGAAATTTGAAAAAGTCCCGATCTCT

                                                        Sph I
                                                          |
     CTTTACCTGAGATGGACTGTTGCTGCACTTTAAACTTTTGCAGGCAGCATGCTGCTTACT
3361 ---------+---------+---------+---------+---------+---------+ 3420
     GAAATGGACTCTACCTGACAACGACGTGAAATTTGAAAACGTCCGTCGTACGACGAATGA

     TTTGTCATGTTATTCACCAAAGAAAATGTATGTATATGTCTTTTTTATACAAATATCCAC
3421 ---------+---------+---------+---------+---------+---------+ 3480
     AAACAGTACAATAAGTGGTTTCTTTTACATACATATACAGAAAAAATATGTTTATAGGTG
```

3'AAAACAGTACAATAAGTGGT5' **PRIMER 9** T_m=52°C

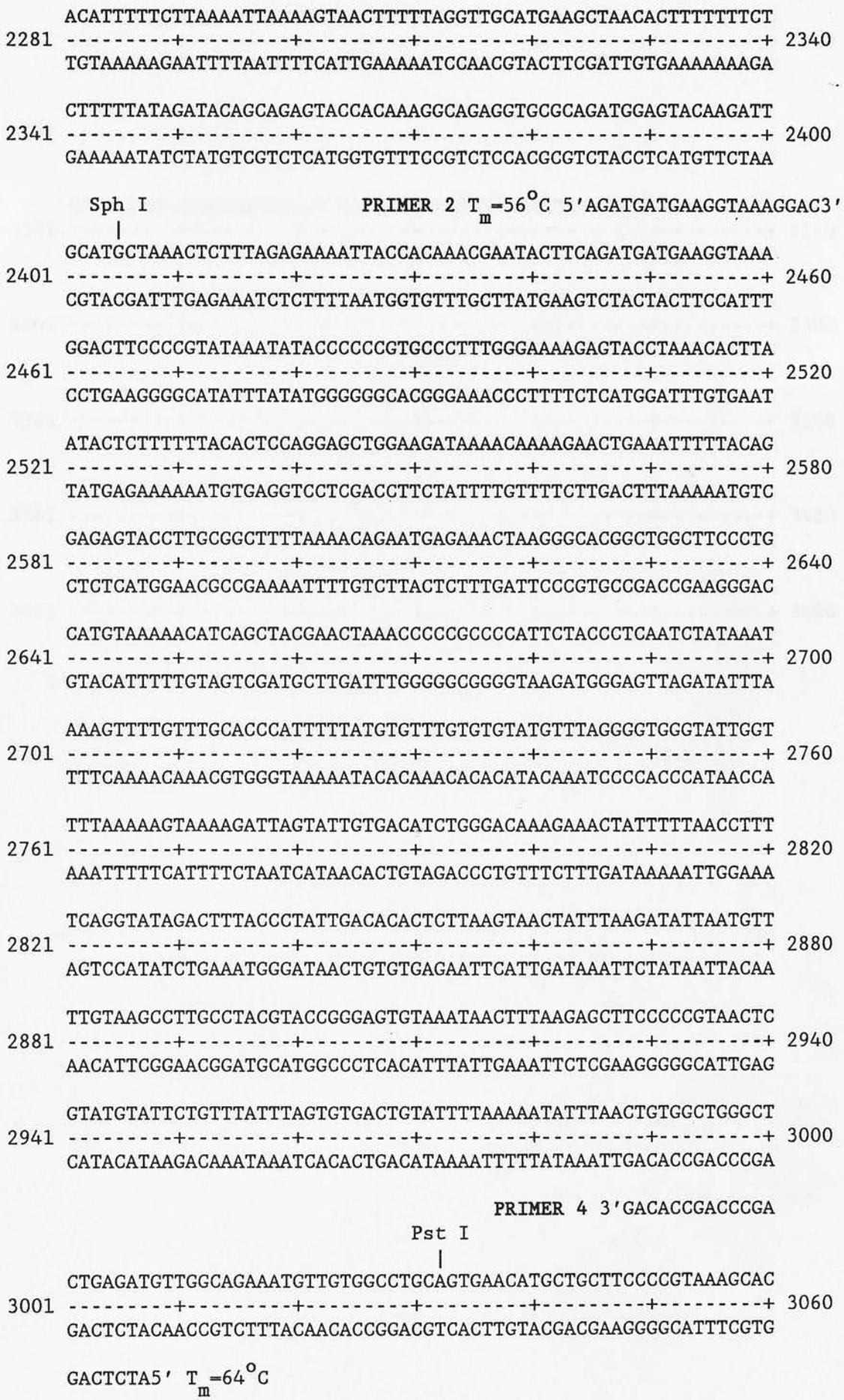
```
     ACATTTTTAAATTACATACTTGTTTCTGTAACATGTGCCTGTAAACTCTGATATAAGCT
3481 ---------+---------+---------+---------+---------+--------- 3539
     TGTAAAAATTTAATGTATGAACAAAGACATTGTACACGGACATTTGAGACTATATTCGA
```

Appendix 3.3 The complete sequence of VIR-2

```
    AGCTTTGAACCAAGGGAGAATGTTTAGCGGAGGGAGACAAAACATGATGAATGCCGCCAA
  1 ---------+---------+---------+---------+---------+---------+ 60
    TCGAAACTTGGTTCCCTCTTACAAATCGCCTCCCTCTGTTTTGTACTACTTACGGCGGTT

    GAAGGTTTTGACAGTGTACTCTAGCCTCAGAGACGATGGAGAAATCAGTCCAGAAGTAAA
 61 ---------+---------+---------+---------+---------+---------+ 120
    CTTCCAAAACTGTCACATGAGATCGGAGTCTCTGCTACCTCTTTAGTCAGGTCTTCATTT

    GGCTTACATGGCTTACATTTTCCAGTGCCTGACATTGAGCAGGTATTCCAACCACTGTTT
121 ---------+---------+---------+---------+---------+---------+ 180
    CCGAATGTACCGAATGTAAAAGGTCACGGACTGTAACTCGTCCATAAGGTTGGTGACAAA

    AAGTTGGAGCAAGAAATCAGAAAAGGTAAAGCCACCTTGACACAGTCTCTGCTGTTTGCA
181 ---------+---------+---------+---------+---------+---------+ 240
    TTCAACCTCGTTCTTTAGTCTTTTCCATTTCGGTGGAACTGTGTCAGAGACGACAAACGT

    CCAAGGAAAGATGTCCCAATAAGACTGTGTTCTCACAGTATGGGGCCAGGCAGAGGTATG
241 ---------+---------+---------+---------+---------+---------+ 300
    GGTTCCTTTCTACAGGGTTATTCTGACACAAGAGTGTCATACCCCGGTCCGTCTCCATAC

                                          Hind II
                                            |
    TGTTACCAGAAGTTTTACTGGAAGCAAGTGAACCACAGTCAACCCTAGCTTACGGGAGCC
301 ---------+---------+---------+---------+---------+---------+ 360
    ACAATGGTCTTCAAAATGACCTTCGTTCACTTGGTGTCAGTTGGGATCGAATGCCCTCGG

                    PRIMER 3 T_m=54°C 5'CTCAAGAAAATACAGACGAA3'

    CTGACATTTCATCTCTTTTGAGAGATTCATCTTCAACTCAAGAAAATACAGACGAAGAAT
361 ---------+---------+---------+---------+---------+---------+ 420
    GACTGTAAAGTAGAGAAAACTCTCTAAGTAGAAGTTGAGTTCTTTTATGTCTGCTTCTTA

              Pst I
                |
    CTGGCCCGTGCTGCAGCAAGACTCTCAGCCCAGGTGTGCCACAGCCTCAATCTGAATATG
421 ---------+---------+---------+---------+---------+---------+ 480
    GACCGGGCACGACGTCGTTCTGAGAGTCGGGTCCACACGGTGTCGGAGTTAGACTTATAC

    ACCCCAGCCCGACCTCTCCACCTGATTCGGACACTGAGTCATGTGATAGCCAGGTTCGTC
481 ---------+---------+---------+---------+---------+---------+ 540
    TGGGGTCGGGCTGGAGAGGTGGACTAAGCCTGTGACTCAGTACACTATCGGTCCAAGCAG

    CAGAAAGTACTGACTCCGACACCCATGCAGAAGACGATGATGTTCCTGAAGCACCCCAAG
541 ---------+---------+---------+---------+---------+---------+ 600
    GTCTTTCATGACTGAGGCTGTGGGTACGTCTTCTGCTACTACAAGGACTTCGTGGGGTTC

                                                    Pst I
                                                      |
    CAGCAAGCCAAACTCAGCCAACAACTACCCAAGAGACACAGAGTTGCTGCAGTGTCTACA
601 ---------+---------+---------+---------+---------+---------+ 660
    GTCGTTCGGTTTGAGTCGGTTGTTGATGGGTTCTCTGTGTCTCAACGACGTCACAGATGT

    ATCCAGCTACCACTCAGACGGGATTCTACTATAACCAGCAGTCAAGTGATAATTTTGCCT
661 ---------+---------+---------+---------+---------+---------+ 720
    TAGGTCGATGGTGAGTCTGCCCTAAGATGATATTGGTCGTCAGTTCACTATTAAAACGGA
```


CTCGTATAGATGCATCTGCCACATCAACCAGTTATGGGTATCCAGGCCCAGTCACTAATC
721
780
GAGCATATCTACGTAGACGGTGTAGTTGGTCAATACCCATAGGTCCGGGTCAGTGATTAG
3'TATCTACGTAGACGGTGTAG5' PRIMER 14 T_m=58°C
ACGGCTTTTCAACTTCTGTCTTTTCCCATACAGTACCCACAACGGCCACTGGCAGCAGCA
781
840
TGCCGAAAAGTTGAAGACAGAAAAGGGTATGTCATGGGTGTTGCCGGTGACCGTCGTCGT
Pvu II
GCTGCATCAAAATATGTATGGGGGTGGTCAGCAGACGACCACATACGGCAGCTATGTGGG
841
900
CGACGTAGTTTTATACATACCCCCACCAGTCGTCTGCTGGTGTATGCCGTCGATACACCC
GGGTTACTCAGATGCTAATGGCCAGTCTGTTGGTGCAAGTACCAGTTACTACACCAGTAA
901
960
CCCAATGAGTCTACGATTACCGGTCAGACAACCACGTTCATGGTCAATGATGTGGTCATT
Pst I
GCCTGCGACCAGCCGCAGCAGCACCGCTCTGCAGTGGACCACTTTGTTTCCACCGGCCAG
961
1020
CGGACGCTGGTCGGCGTCGTCGTGGCGAGACGTCACCTGGTGAAACAAAGGTGGCCGGTC
CTCATCTGCTTCCAGCAGCCAAGTTTCTGATTATTTCAGCCCCTTCCCTTATTTTCCTGG
1021
1080
GAGTAGACGAAGGTCGTCGGTTCAAAGACTAATAAAGTCGGGGAAGGGAATAAAAGGACC
PRIMER 8 3'TTCAAAGACTAATAAAGTCG5' T_m=52°C
CAGTTCTACAGTTCCTCAGGCCTTCGAGCCTTTAGCGCCATCTACCCCTAGCTTGCTGGA
1081
1140
GTCAAGATGTCAAGGAGTCCGGAAGCTCGGAAATCGCGGTAGATGGGGATCGAACGACCT
TGAACTGCTTGATAGAGACAGTGGACTTGTAAGCCAGCAGCAGGCACCTGCACCTCCCCA
1141
1200
ACTTGACGAACTATCTCTGTCACCTGAACATTCGGTCGTCGTCCGTGGACGTGGAGGGGT
AAATGACCAGGGGGGGCCTCCTCAGTATGTGCCCGTTGCTCAAGAGCAACAGCAGTCCAG
1201
1260
TTTACTGGTCCCCCCCGGAGGAGTCATACACGGGCAACGAGTTCTCGTTGTCGTCAGGTC
TACAGACCCGCTGTCTGATGAGATGAGGAGAATTTTTTGAGTTTTTGACAGTGTGAACCC
1261
1320
ATGTCTGGGCGACAGACTACTCTACTCCTCTTAAAAAAACTCAAAAACTGTCACACTTGGG
CGTTTGAACCAGACCATAGCCATATTTTGGGAGTGTAAACAGAATCCTCTGTAAACTGCC
1321
1380
GCAAACTTGGTCTGGTATCGGTATAAAACCCTCACATTTGTCTTAGGAGACATTTGACGG
TTTTTTATGTGAAATGCGTCTCTCGCCAAAACTATCTGCTCTTGTTTTGACAAAAAGTGG
1381
1440
AAAAAATACACTTTACGCAGAGAGCGGTTTTGATAGACGAGAACAAAACTGTTTTTCACC
AAGTGTGTGTTTATATGATTCACAAAACTGTAGTTTAACAGTAAAAGTTAATTGTGGTTT
1441
1500
TTCACACACAAATATACTAAGTGTTTTGACATCAAATTGTCATTTTCAATTAACACCAAA

```
     GTGGTATGACACACAGTTCTGGAAACTGACCAAACAGTGTGTTCTTATGTGCACTTATTT
1501 ---------+---------+---------+---------+---------+---------+ 1560
     CACCATACTGTGTGTCAAGACCTTTGACTGGTTTGTCACACAAGAATACACGTGAATAAA

     TGTATTTCCCTTATGCCTGCCAGAGTGCTCAATAAAAGTTACACTCAAGTGCCACTGAGA
1561 ---------+---------+---------+---------+---------+---------+ 1620
     ACATAAAGGGAATACGGACGGTCTCACGAGTTATTTTCAATGTGAGTTCACGGTGACTCT

       PRIMER 1 5'TCAACACTATAAAAGCATCT3' T_m=52°C

     GTTTTTTCTTTTTTCAACACTATAAAAGCATCTTTGGGGACTCATGCATAAGCACTCTGC
1621 ---------+---------+---------+---------+---------+---------+ 1680
     CAAAAAAGAAAAAAGTTGTGATATTTTCGTAGAAACCCCTGAGTACGTATTCGTGAGACG

       Pvu II
         |
     TACAGCTGACATGCTACCCATTTCAGCATTTAGCCCTCTGAAATTCTTGGGTGGAGTTGGT
1681 ---------+---------+---------+---------+---------+---------+ 1740
     ATGTCGACTGTACGATGGGTAAAGTCGTAAATCGGGAGACTTTAAGAACCCACCTCAACCA

     GGCTACACCCCAGGGGCCAGTATTGTGGGATGCCACTACTACTCCTGTCCCTTTCCAGGC
1741 ---------+---------+---------+---------+---------+---------+ 1800
     CCGATGTGGGGTCCCCGGTCATAACACCCTACGGTGATGATGAGGACAGGGAAAGGTCCG

     TTTCCCCTGCCCTCCGGCAATAACTCCTATAAATCAGTTTGTTAATTATCAGCCTGCGCC
1801 ---------+---------+---------+---------+---------+---------+ 1860
     AAAGGGGACGGGAGGCCGTTATTGAGGATATTTAGTCAAACAATTAATAGTCGGACGCGG

     TGCGATGGATAAGATATGGAGACCTGCCTTTGAGGACTACCGTGGGGTAACTATGAGACC
1861 ---------+---------+---------+---------+---------+---------+ 1920
     ACGCTACCTATTCTATACCTCTGGACGGAAACTCCTGATGGCACCCCATTGATACTCTGG

                 Kpn I
                   |
     GGCCTTACCGGAGGTACCTGAAACAAGTTTGAGACCAGCTTTACAGGAGCTACCCGAGCC
1921 ---------+---------+---------+---------+---------+---------+ 1980
     CCGGAATGGCCTCCATGGACTTTGTTCAAACTCTGGTCGAAATGTCCTCGATGGGCTCGG

     ATCATCTCCTCAAAGTCAGAGTTCTGTGGATGATGACACAGATTCTAAAGAAGATGTTAC
1981 ---------+---------+---------+---------+---------+---------+ 2040
     TAGTAGAGGAGTTTCAGTCTCAAGACACCTACTACTGTGTCTAAGATTTCTTCTACAATG

     AGAGACCTTAGAATGTGCTCAGGCTCTGACAGATCTGAAGTGGGGTACAGTGGACCCTTC
2041 ---------+---------+---------+---------+---------+---------+ 2100
     TCTCTGGAATCTTACACGAGTCCGAGACTGTCTAGACTTCACCCCATGTCACCTGGGAAG

     CGAGAAGACATCTCAGTTACTGCTTGCACTTCAAGCGGCGAGACCAGCATTTGCAGAAGG
2101 ---------+---------+---------+---------+---------+---------+ 2160
     GCTCTTCTGTAGAGTCAATGACGAACGTGAAGTTCGCCGCTCTGGTCGTAAACGTCTTCC

     CAAAGGAACAGCTATACGGGGCTACAGACTTATTGAAGCTAAAGACTTGGGCAAGCTAGT
2161 ---------+---------+---------+---------+---------+---------+ 2220
     GTTTCCTTGTCGATATGCCCCGATGTCTGAATAACTTCGATTTCTGAACCCGTTCGATCA

                                                        Kpn I
                                                          |
     AGTGAGAGGACCCAAAAAAGACAAGCGAGATCCAGACTTCTACATCAAAAAGGTACCTTT
2221 ---------+---------+---------+---------+---------+---------+ 2280
     TCACTCTCCTGGGTTTTTTCTGTTCGCTCTAGGTCTGAAGATGTAGTTTTTCCATGGAAA
```


ACATTTTTCTTAAAATTAAAAGTAACTTTTTAGGTTGCATGAAGCTAACACTTTTTTTCT
2281
2340
TGTAAAAAGAATTTTAATTTTCATTGAAAAATCCAACGTACTTCGATTGTGAAAAAAAGA
CTTTTTATAGATACAGCAGAGTACCACAAAGGCAGAGGTGCGCAGATGGAGTACAAGATT
2341
2400
GAAAAATATCTATGTCGTCTCATGGTGTTTCCGTCTCCACGCGTCTACCTCATGTTCTAA
Sph I
PRIMER 2 T_m=56°C 5'AGATGATGAAGGTAAAGGAC3'
GCATGCTAAACTCTTTAGAGAAAATTACCACAAACGAATACTTCAGATGATGAAGGTAAA
2401
2460
CGTACGATTTGAGAAATCTCTTTTAATGGTGTTTGCTTATGAAGTCTACTACTTCCATTT
GGACTTCCCCGTATAAATATACCCCCCGTGCCCTTTGGGAAAAGAGTACCTAAACACTTA
2461
2520
CCTGAAGGGGCATATTTATATGGGGGGCACGGGAAACCCTTTTCTCATGGATTTGTGAAT
ATACTCTTTTTTACACTCCAGGAGCTGGAAGATAAAACAAAAGAACTGAAATTTTTACAG
2521
2580
TATGAGAAAAAATGTGAGGTCCTCGACCTTCTATTTTGTTTTCTTGACTTTAAAAATGTC
GAGAGTACCTTGCGGCTTTTAAAACAGAATGAGAAACTAAGGGCACGGCTGGCTTCCCTG
2581
2640
CTCTCATGGAACGCCGAAAATTTTGTCTTACTCTTTGATTCCCGTGCCGACCGAAGGGAC
CATGTAAAAACATCAGCTACGAACTAAACCCCCGCCCCATTCTACCCTCAATCTATAAAT
2641
2700
GTACATTTTTGTAGTCGATGCTTGATTTGGGGGCGGGGTAAGATGGGAGTTAGATATTTA
AAAGTTTTGTTTGCACCCATTTTTATGTGTTTGTGTGTATGTTTAGGGGTGGGTATTGGT
2701
2760
TTTCAAAACAAACGTGGGTAAAAATACACAAACACACATACAAATCCCCACCCATAACCA
TTTAAAAAGTAAAAGATTAGTATTGTGACATCTGGGACAAAGAAACTATTTTTAACCTTT
2761
2820
AAATTTTTCATTTTCTAATCATAACACTGTAGACCCTGTTTCTTTGATAAAAATTGGAAA
TCAGGTATAGACTTTACCCTATTGACACACTCTTAAGTAACTATTTAAGATATTAATGTT
2821
2880
AGTCCATATCTGAAATGGGATAACTGTGTGAGAATTCATTGATAAATTCTATAATTACAA
TTGTAAGCCTTGCCTACGTACCGGGAGTGTAAATAACTTTAAGAGCTTCCCCCGTAACTC
2881
2940
AACATTCGGAACGGATGCATGGCCCTCACATTTATTGAAATTCTCGAAGGGGGCATTGAG
GTATGTATTCTGTTTATTTAGTGTGACTGTATTTTAAAAATATTTAACTGTGGCTGGGCT
2941
3000
CATACATAAGACAAATAAATCACACTGACATAAAATTTTTATAAATTGACACCGACCCGA
PRIMER 4 3'GACACCGACCCGA
Pst I
CTGAGATGTTGGCAGAAATGTTGTGGCCTGCAGTGAACATGCTGCTTCCCCGTAAAGCAC
3001
3060
GACTCTACAACCGTCTTTACAACACCGGACGTCACTTGTACGACGAAGGGGCATTTCGTG
GACTCTA5' T_m=64°C

```
     TTTTGGTAGACATTTTTTTTATTTTGGCAGCCACAAATTTAATGATAGCTGCCTTTGCCC
3061 ---------+---------+---------+---------+---------+---------+ 3120
     AAAACCATCTGTAAAAAAAATAAAACCGTCGGTGTTTAAATTACTATCGACGGAAACGGG

     TGGGCTGCTTGGCTTTCTATAAGCAATTGGTTTATATAACAATTGGCAACCTAACCTTTC
3121 ---------+---------+---------+---------+---------+---------+ 3180
     ACCCGACGAACCGAAAGATATTCGTTAACCAAATATATTGTTAACCGTTGGATTGGAAAG

     CCCACCAGAGCGGGGATGAGGTTATTAGGGCCATGTACATACCCCCAGTCAATGACAGTG
3181 ---------+---------+---------+---------+---------+---------+ 3240
     GGGTGGTCTCGCCCCTACTCCAATAATCCCGGTACATGTATGGGGGTCAGTTACTGTCAC

     TGGATTTTAACCCTGGGTTCAGACTTAGCTGGTTAAACACCCTAAGTCCGCTCTCCGATG
3241 ---------+---------+---------+---------+---------+---------+ 3300
     ACCTAAAATTGGGACCCAAGTCTGAATCGACCAATTTGTGGGATTCAGGCGAGAGGCTAC

     GGCCCTATGACTCTTGGTCCCAGTGCGAGATCTGCCCCGGTCGCTTTGTGTCCAAAAGGC
3301 ---------+---------+---------+---------+---------+---------+ 3360
     CCGGGATACTGAGAACCAGGGTCACGCTCTAGACGGGGCCAGCGAAACACAGGTTTTCCG

     CTGTTATTATGTCCCCCCAAAGACCTATAGCTTTCAAAACTGCTTCTTTGCCTGTAAGAA
3361 ---------+---------+---------+---------+---------+---------+ 3420
     GACAATAATACAGGGGGGTTTCTGGATATCGAAAGTTTTGACGAAGAAACGGACATTCTT

     TATTTCTAAATGTTTTTATCTTTATACTCCACAGAATATAACTGATCCTTTTTTTGACCA
3421 ---------+---------+---------+---------+---------+---------+ 3480
     ATAAAGATTTACAAAAATAGAAATATGAGGTGTCTTATATTGACTAGGAAAAAAACTGGT
```

5'CACCTTACGAGACCAGGAT3' **PRIMER 16** T_m=60°C

```
     CACCTTACGAGACCAGGATATCTGGATAGGGACTTTTTTTAAAAAGCT
3481 ---------+---------+---------+---------+-------- 3528
     GTGGAATGCTCTGGTCCTATAGACCTATCCCTGAAAAAAATTTTTCGA
```